Příloha 4 – Plný text manuskriptu: Hojny J, Zemankova P, Lhota F, Sevcik J, Stranecky V, Hartmannova H, Hodanova K, Mestak O, Pavlista D, Janatova M, Soukupova J, Vocka M, Kleibl Z, Kleiblova P. **Multiplex PCR and NGS-based identification of mRNA splicing variants: Analysis of BRCA1 splicing pattern as a model.** *Gene.* 2017 Dec 30;637:41-49. doi:10.1016/j.gene.2017.09.025.

Gene 637 (2017) 41–49

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Methodological paper

Multiplex PCR and NGS-based identification of mRNA splicing variants: Analysis of BRCA1 splicing pattern as a model

Jan Hojny[a], Petra Zemankova[a], Filip Lhota[a], Jan Sevcik[a], Viktor Stranecky[b], Hana Hartmannova[b], Katerina Hodanova[b], Ondrej Mestak[c], David Pavlista[d], Marketa Janatova[a], Jana Soukupova[a], Michal Vocka[e], Zdenek Kleibl[a], Petra Kleiblova[a,f,*]

[a] Institute of Biochemistry and Experimental Oncology, First Faculty of Medicine, Charles University, Prague 12853, Czech Republic
[b] Institute of Inherited Metabolic Disorders, First Faculty of Medicine, Charles University and General University Hospital in Prague, Prague 120 00, Czech Republic
[c] Department of Plastic Surgery, First Faculty of Medicine, Charles University and Na Bulovce Hospital, Prague 180 81, Czech Republic
[d] Department of Obstetrics and Gynecology, First Faculty of Medicine, Charles University and General University Hospital in Prague, Prague 120 00, Czech Republic
[e] Department of Oncology, First Faculty of Medicine, Charles University and General University Hospital in Prague, Prague 120 00, Czech Republic
[f] Institute of Biology and Medical Genetics, First Faculty of Medicine, Charles University and General University Hospital in Prague, Prague 120 00, Czech Republic

ARTICLE INFO

Keywords:
Alternative splicing
mRNA splicing variant
BRCA1
Multiplex PCR
NGS

ABSTRACT

Alternative pre-mRNA splicing increases transcriptome plasticity by forming naturally-occurring alternative splicing variants (ASVs). Alterations of splicing processes, caused by DNA mutations, result in aberrant splicing and the formation of aberrant mRNA isoforms. Analyses of hereditary cancer predisposition genes reveal many DNA variants with unknown clinical significance (VUS) that potentially affect pre-mRNA splicing. Therefore, a comprehensive description of ASVs is an essential prerequisite for the interpretation of germline VUS in high-risk individuals.

To identify ASVs in a gene of interest, we have proposed an approach based on multiplex PCR (mPCR) amplification of all theoretically possible exon-exon junctions and subsequent characterization of size-selected and pooled mPCR products by next-generation sequencing (NGS). The efficiency of this method is illustrated by a comprehensive analysis of *BRCA1* ASVs in human leukocytes, normal mammary, and adipose tissues and stable cell lines.

We revealed 94 BRCA1 ASVs, including 29 variants present in all tested samples. While differences in the qualitative expression of BRCA1 ASVs among the analyzed human tissues were minor, larger differences were detected between tissue and cell line samples.

Compared with other ASV analysis methods, this approach represents a highly sensitive and rapid alternative for the identification of ASVs in any gene of interest.

1. Introduction

Hereditary mutations in cancer-susceptibility genes are responsible for tumor development in about 5% of all cancer patients. The carriers of hereditary mutations face a high life-time risk of cancer, which often develops at an early age (Rahman, 2014). Tailored care improving life expectancy in these high-risk individuals requires an unequivocal identification of causative mutations in hundreds of known cancer-susceptibility genes. The recent introduction of next-generation sequencing (NGS) into clinical diagnostics enables a simultaneous analysis of multiple genes; however, its clinical utility is hampered by the presence of many variants with unknown significance (VUS) (Cheon et al., 2014). These genetic changes emerge as rare germline missense, silent or intronic variants with an uncertain biological and functional impact on the resulting protein isoform. The number of identified VUS rises proportionally to the length of the analyzed genomic sequence, and many of them may alter mRNA splicing processes (Tavtigian & Chenevix-Trench, 2014).

Pre-mRNA splicing controls the composition of matured mRNA by regulated intron exclusion and exon linking. A primary wild-type (wt) transcript (pre-mRNA) can be variably processed by alternative splicing into alternative mRNA variants translated into protein isoforms with

Abbreviations: ASV(s), alternative splicing variant(s); mPCR, multiplex PCR; NGS, next generation sequencing; SAS, splice acceptor shift; SDS, splice donor shift; VUS, variant of unknown clinical significance

* Corresponding author at: Institute of Biochemistry and Experimental Oncology, First Faculty of Medicine, Charles University, U Nemocnice 5, Prague 2 128 53, Czech Republic.
E-mail address: petra.kleiblova@lf1.cuni.cz (P. Kleiblova).

http://dx.doi.org/10.1016/j.gene.2017.09.025
Received 28 June 2017; Received in revised form 26 July 2017; Accepted 13 September 2017
Available online 14 September 2017
0378-1119/

different biological activities (Bentley, 2014). Alternative splicing must be distinguished from aberrant splicing resulting from a dysregulation of natural splice site recognition caused by DNA mutations. The DNA sequence variants affecting pre-mRNA splicing are more prevalent than estimated up to now, and they account for at least 15% of disease-causing mutations, and for up to 50% of all mutations described in some genes (Caminsky et al., 2015; Soukarieh et al., 2016). Various splicing assays help to disclose the impact of VUS on splicing processes (Whiley et al., 2014), and variants that cause aberrant splicing are considered pathogenic. The evaluation of aberrant splicing requires a precise knowledge of alternative splicing variants (ASVs) for the analyzed primary transcript (Colombo et al., 2014). Despite large-scale RNA sequencing projects (e.g. ENCODE, GTEx), there is no precise catalogue of ASVs or validated RNAseq data for most clinically-relevant genes (Sloan et al., 2016; Baralle & Buratti, 2017).

Recently, two articles have described a comprehensive analysis of naturally occurring splicing variants in *BRCA1* (Colombo et al., 2014; Romero et al., 2015), one of the most studied cancer-susceptibility genes responsible for hereditary breast and ovarian cancer (Kleibl & Kristensen, 2016). In both studies, the RNA-based analysis involved a combination of various techniques including RT-PCR, exon scanning, cloning, sequencing, and relative (semi)quantification. This experimental variability negatively affects the reproducibility of splicing variant analyses from various mRNA sources and makes the methods difficult to use in the analyses of other gene products. The large number of these analytic techniques makes the analysis laborious, may negatively affects its reproducibility, and adaptation to characterize another gene transcripts. Therefore, we aimed to develop a versatile approach suitable for the characterization of ASVs in any gene of interest based on NGS of multiplex PCR-generated amplicons covering all theoretically possible exon-exon junctions.

2. Materials and methods

2.1. An overview of experimental design

We aim to characterize the ASVs of any gene from an RNA sample, dominantly on the qualitative level. The analysis comprises four steps: i) multiplex PCR (mPCR) amplification of all theoretically possible mRNA splicing variants from a cDNA template, ii) pooling of mPCRs, purification and size selection of pooled mPCR products targeting short amplicons, iii) standard NGS library preparation from size-selected mPCR fragments followed by routine Illumina sequencing, and iv) a bioinformatics analysis. We have demonstrated the efficiency of the mPCR/NGS approach by the characterization of BRCA1 ASVs because i) the *BRCA1* gene is the most frequently altered breast cancer susceptibility gene in many countries including the Czech Republic and many BRCA1 VUS contribute to aberrant splicing, ii) 63 BRCA1 mRNA variants were recently described using conventional RT-PCR and capillary-electrophoresis by Colombo et al. (2014) and Romero et al. (2015), indicating that iii) the BRCA1 mRNA splicing isoform pattern is highly variable.

2.2. Alternative BRCA1 splice site nomenclature

All alternative splicing events were classified into biotypes based on previously published nomenclature (Colombo et al., 2014; Romero et al., 2015). The insertions (▼) and deletions (Δ) denote splicing events affecting a single exon (cassette) or > 1 consecutive exons (multicassette). The deletions affecting the 5′ and 3′ ends of an exon were described as an exon number with an added “p” or “q”, respectively. The extension of an exon sequence into an adjacent intronic region is described as an exon number with an “a”. The splice donor/acceptor shift (SDS/SAS) variants were identified and counted as NGS reads with deletions of nucleotides at the exon-exon junctions or insertions of intronic parts flanking to the 5′ or 3′ ends of an exon. The mixed biotypes denoted combinations of the above-mentioned events. The *BRCA1* exons were numbered according to the Breast Cancer Information Core Database (https://research.nhgri.nih.gov/bic/) nomenclature.

2.3. Patients and samples

The characterization of BRCA1 ASVs was performed in 96 RNA samples obtained from 32 individuals (Supplementary Table S1), including 16 non-cancer controls, eight breast-cancer (BC) patients without *BRCA1* mutation, and eight *BRCA1*-mutation carriers. Simultaneously-obtained tissue samples were collected during BC surgery or preventive mastectomy (in BC patients and *BRCA1* mutation carriers) or during cosmetic breast surgery (in controls). All enrolled individuals were Caucasians of a Czech origin who gave a written informed consent approved by ethical committees to participate in the study. RNA samples were isolated from the leukocytes and macroscopically dissected fresh mammary and adipose perimammary tissues of each individual. We further analyzed RNA samples from stable human cell lines (from MCF7 cells, and from pooled EM-G3, HeLa, and MDA-MB-231 cells). The cell lines were maintained as described previously (Brozova et al., 2007; Sevcik et al., 2012; Sevcik et al., 2013; Vondruskova et al., 2008).

2.3.1. Total RNA isolation, quality control and cDNA synthesis

All RNA samples were processed according to MIQE guidelines (Bustin et al., 2009). Peripheral blood samples (2.5 ml) were collected into PAXgene Blood RNA tubes, incubated overnight at room temperature, and stored at − 20 °C. All stored samples were thawed and stored for 2 h at room temperature before RNA isolation performed with PAXgene Blood RNA Kit (PreAnalytiX). Solid tissue samples (~100 mg/sample) were submerged into 1 ml of RNAlater (Qiagen) immediately after surgical excision, processed according to the manufacturer, and after overnight incubation (2–8 °C) stored at − 80 °C until RNA isolation. Forty micrograms of thawed, RNAlater-preserved samples were homogenized using MagNA Lyser Green Beads tubes on MagNA Lyser Instrument (Roche) in the presence of 1 ml Qiazol (Qiagen). Total RNAs from the homogenated tissues and cultured cells were isolated with RNeasy Tissue Mini Kit (Qiagen).

All RNA samples were treated by DNase I, quantified on NanoDrop 1000 (Thermo Fisher Scientific) and characterized by the RNA integrity number (RIN) using Bioanalyzer 2100 with RNA 6000 Nano Kit (Agilent Technologies; tissue samples RIN_{mean} = 7.4; range 6.3–8.9).

Overall, 1.5 μg of RNA was used for cDNA synthesis (in a reaction volume of 20 μl). The cDNA synthesis was performed using SuperScript III Reverse Transcriptase (Thermo Fisher Scientific) and random hexamers (Roche) as described previously (Kleiblova et al., 2010). A routine PCR control of cDNA quality/integrity was performed prior to further analyses (not shown).

2.4. Multiplex PCR (mPCR) amplification and size selection

2.4.1. Primer designing

The primers were designed to specifically cover all possible exon-exon junctions. The resulting PCR amplicons thus enable the identification of all canonical as well as alternative splicing mRNA isoforms. Forward primers targeted the 3′ region while reverse primers aimed at the 5′ region of an exon (Supplementary Fig. S1). For the analysis of a single gene transcript consisting of *N* exons, the number of *N*-2 forward and *N*-2 reverse primers is required for the amplification of all theoretically possible exon-exon junctions in at least *N*-2 mPCR reactions that are finally pooled into one mPCR pool.

For the analysis of BRCA1 ASVs, we designed 45 primers targeting 22 coding exons of the canonical BRCA1 transcript (NM_007294), and alternative exons 11q and 13A (Fig. 1A, Supplementary Table S2). All individual PCRs were optimized separately (Supplementary Fig. S2)

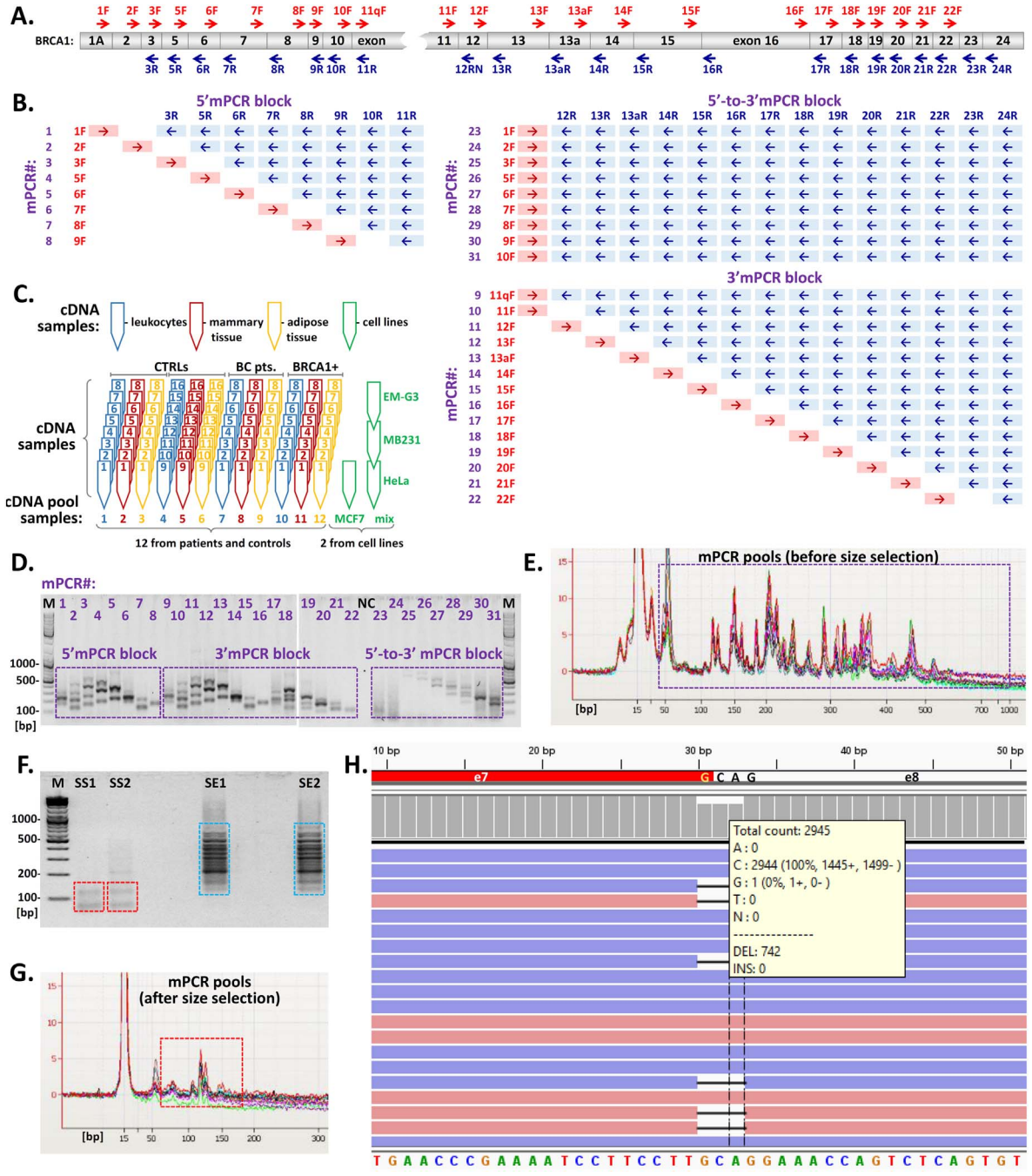


Fig. 1. Method overview. The chart (A) shows primer pairs (forward – red; reverse – blue) used for mPCR amplifications of all theoretically possible BRCA1 ASVs. Due to the presence of the large exon 11, 31 mPCR reactions (violet letters; B) were performed in three 'blocks'. All 31 mPCR reactions were performed with each of 14 cDNA pools (C). Twelve cDNA pools of human cDNA samples and two cDNA pools from cell line samples served as templates for 31 mPCR reactions. Agarose gel electrophoresis of 31 mPCRs (ranging in size between ~50 and ~700 bp; purple dashed line) from a single cDNA pool is shown in (D). All 31 mPCRs from each cDNA pool sample were further pooled together (to create an mPCR pool) and analyzed on Agilent Bioanalyzer (E; the overlaying electrophoretograms of 12 mPCR pools show good reproducibility). The double-sided size selection was used to enrich the short mPCR amplicons that were subsequently used for NGS library preparation. The agarose gel electrophoresis (F) displays the enrichment of size-selected fragments (SS; red dashed line boxes) while size-excluded fragments (SE; blue dashed line boxes) were discarded. The size-selected (SS) samples (ranging at 50–150 bp in length; red dashed line box in G) were verified by Agilent Bioanalyzer. The MiSeq reads were mapped to the bam files (Supplementary Table S3) containing all theoretically possible BRCA1 splicing cassette and multicassette events. Visual inspection of reads in IGV viewer enabled direct assessment and quantification of SDS/SAS as shown (H) for the ΔCAG at the 5′ end of exon 8 (erroneously mapped as ΔGCA; coverage depth is shown as grey vertical bars; forward (pink) and reverse (blue) reads are shown as vertical bars). In normal mammary tissue of BC patients (shown in H); ΔCAG accounted for 742 reads (20.1%), while 2945 sequencing reads were recorded for wt sequence (Supplementary Table S6). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

and subsequently in mPCR (Supplementary Fig. S3).

2.4.2. *mPCR amplification and preparation of mPCR pools*

All theoretically possible splicing events of a gene of interest can be amplified in a few mPCRs. The number of mPCRs depends primarily on the number of exons of the analyzed gene. Every individual mPCR contains a single forward primer and a set of reverse primers targeting all consecutive exons except the exon directly flanking to the exon targeted by the forward primer (Supplementary Fig. S1). The undesired synthesis of the amplicons of several consecutive exons was reduced by

a short elongation time (Supplementary Fig. S3B).

The analysis of BRCA1 ASVs in each cDNA template required 31 mPCRs (Fig. 1B). To overcome the usually limited amount of RNA (and resulting cDNA) available from human tissue samples, the cDNA template for mPCRs was prepared by the pooling of individual cDNAs (from the same tissue and sample subgroup; Fig. 1C). This may be necessary especially with multiexonic and low-expressed genes (including *BRCA1*) which require a larger number of mPCRs consuming an increased amount of cDNA. Although the cDNA pooling resulted in a loss of information about the expression of ASVs in the individual cDNA samples, it increased the chances of detecting low-expressed ASVs.

For the BRCA1 analysis, eight individual cDNA samples (16 μl each) from the same tissue type and patient group were pooled together to obtain 12 patient cDNA pools (each 128 μl; Fig. 1C). We also analyzed cDNAs from stable human cell lines (cDNA from MCF7 cells and a pool of cDNAs from EM-G3, HeLa and MD-MB-231). Each of the 12 patient cDNA pools, MCF7 cDNA, and cDNA cell line pools served as a template for 31 mPCRs. Each 40 μl mPCR contained 4 μl of pooled cDNA template (equivalent to 300 ng of RNA), a single forward primer (final concentration 225 nM), a variable set of reverse primers (final concentration 75 nM of each), and FastStart Taq DNA Polymerase (Roche) according to the manufacturer's instructions. The mPCR amplifications involved 4-minute incubation at 95 °C followed by 35 cycles (95 °C for 10 s, 62 °C for 20 s, and 72 °C for 15 s) and final extension at 72 °C for 7 min. The individual mPCR products were analyzed electrophoretically (Fig. 1D). After that, 35 μl of each of the 31 mPCRs from a single cDNA pool were mixed together to provide an mPCR pool. The resulting mPCR pools were characterized by capillary electrophoresis using 2100 Bioanalyzer and DNA 1000 Kit (Agilent Technologies; Fig. 1E).

2.4.3. Size selection and purification of mPCR pools

To reduce the presence of longer mPCR-amplified fragments containing short consecutive exons, the mPCRs pools were subjected to size selection using double-sided solid phase reversible immobilization with magnetic beads. The size-selected and purified amplicons served as templates to prepare a standard NGS library.

As the length of the targeted mPCR amplicons of BRCA1 ASVs was expected to range mostly at 80–90 bp, we performed size selection using magnetic beads to remove undesired amplicons (< 50 bp and > 150 bp; Fig. 1F). First, we used 1.8 × concentration of Agencourt AMPure XP reagent (Beckman Coulter) to bind and remove amplicons > 150 bp (dominantly containing PCR products amplified from the canonical BRCA1 mRNA). Subsequently, we mixed the supernatant from the first reaction with a reagent to the final 2.5 × concentration to withdraw DNA fragments > 50 bp in length. The size-selected and purified samples were characterized on Agilent Bioanalyzer with DNA 1000 Kit (Fig. 1G).

2.4.4. NGS library preparation, MiSeq sequencing

First, Dynazyme II DNA Polymerase (Thermo Fisher Scientific) was used to create 3′-dA overhangs in DNA fragments in size-selected and purified mPCR pools. Subsequently, Illumina sequencing adaptors were ligated using Rapid DNA Ligation Kit (Thermo Fisher Scientific). Finally, the seven-cycle PCR reaction (NEBNext High Fidelity PCR Master Mix, NEB) with a universal primer introduced a 6-bp index sequence unique for each mPCR pool. The processed samples were purified using Agencourt AMPure XP Reagent after each step of sequencing library preparation. The quality of the prepared libraries was characterized on 2100 Bioanalyzer and quantified fluorimetrically (Qubit; Thermo Fisher Scientific). To achieve sufficient sequencing coverage and sample diversity, we admixed our mPCR-prepared libraries with panel sequencing libraries of high complexity to standard MiSeq runs (one mPCR pool library represented 1/30 of sequencing capacity). Runs were sequenced with MiSeq Reagent Kit v3 (150 cycle).

2.4.5. Bioinformatics

Bioinformatics requires double-step mapping. First, sequencing data are mapped to a user-designed fasta file containing the exon-exon junctions of all theoretically possible splicing variants of the gene of interest. Thus we can identify all cassette and multicassette splicing deletions and variants resulting from short SDS/SAS. Second, sequencing data are mapped to the genomic sequence of the analyzed gene in order to identify the exonized intronic sequences.

In our BRCA1 analysis, the primary raw data sets (in the fastq.gz format) were processed by a routine bioinformatics pipeline using software tools specified below using the default settings (if not otherwise specified in the list of commands listed in Supplementary Table S7). The remaining sequences were trimmed (to remove adapters and low-quality bases) before mapping by Trimmomatic (ver. 0.32; http://www.usadellab.org). First, we mapped raw data sets to the prepared fasta file using Novoalign (ver. 2.08.03; Novocraft). Reads with insufficient sequencing quality were removed. This “BRCA1 splicing” fasta file contained 311 sequences (“exon-exon_computed” in Supplementary Table S3) which considered all possible combinations of known BRCA1 exons (NM_007294), including alternative exon 13A, and known SDS/SAS combinations > 10 bp (11q, and 5q). Output in the SAM format was transformed to BAM by Picard tools (ver. 1.129; https://broadinstitute.github.io/picard/) and displayed in IGV (Integrative Genomics Viewer, Broad Institute; Fig. 1H). Coverage statistics were created by SAMtools (ver. 0.1.19; http://samtools.sourceforge.net/). Since this approach ignored the presence of exonized intron sequences, we further mapped all raw data to the *BRCA1* gene sequence (81,189 bp from NG_005905 spanning sequence 92,500–173,688). The mapping results were analyzed in IGV manually to remove reads with soft-clipped bases. Mapped reads that exceeded from exons into flanking introns or reads mapped to deep intronic sequences, respectively, were recorded including sequences of unmapped nucleotides. The unmapped parts of reads were BLASTed with the *BRCA1* gene sequence (https://blast.ncbi.nlm.nih.gov/). If the unmapped sequence contained the 5′ or 3′ parts of a *BRCA1* exon, we were able to identify the entire intronic insertion. Sequences comprising the identified intronic insertions with flanking exonic sequences were added to a fasta file (“from_IGV_BLASTed” in Supplementary Table S3) and used for new mapping from the original dataset. Mapped reads were manually inspected in IGV in order to eliminate incorrectly mapped reads with soft-clipped bases.

To compare the numbers of ASV reads among the examined mPCR pools, we expressed the number of sequencing reads of an ASV as value normalized to 10^6 reads in the given mPCR pool (Supplementary Tables S4 and S5).

3. Results

In order to prepare a versatile method for a direct assessment of splice junction events, we designed the mPCR/NGS-based approach enabling the identification of ASVs in a gene of interest. Using the described method, we identified 94 BRCA1 ASVs (Table 1) comprising all previously described biotypes (Colombo et al., 2014). Our analysis of simultaneously-obtained tissue RNA samples revealed that the highest number of ASVs were expressed in mammary tissue (72 variants), followed by leukocytes and perimammary adipose tissues (67 and 54 variants, respectively).

Forty-eight ASVs were identified in all examined human tissue types, with 29 of them expressed in each cDNA tissue sample (referred here as “ubiquitous”; Fig. 2; Table 1).

Only slight qualitative differences were identified among tissue samples from non-cancer controls, BC patients and *BRCA1*-mutation carriers. In contrast, the spectrum of ASVs differed between tissue and cell-line samples (Fig. 2). Altogether, 76 ASVs (11 of them exclusively) were expressed in cell lines.

The detected variants included 25 in-frame ASVs that may

Table 1
Description of all BRCA1 ASVs identified in this study, including variant name, systematic description of the variant at the cDNA level, functional annotation, and biotype class. The values for the expression of particular variants in analyzed cell lines and human tissue samples show normalized sequencing reads (reads per 10^6 reads), the colours indicate normalized sequencing coverage: 0 reads = white; 1–9 reads = green; 10–99 reads = yellow; 100–999 reads = light red; 1000–9999 reads = red; > 9999 reads = dark red. The "ubiquitous" variants (expressed in all analyzed human tissue samples) are highlighted by bold letters and blue lines. The presence of variants in the analyzed human tissue samples and cell lines was compared to that reported as predominant (P), present (1), or absent (0) in studies by Colombo et al. (2014), Romero et al. (2015) and Orban and Olah (2003). The extended version of this table is shown in Supplementary Table S5.

Variant description	HGVS description	Functional annotation	Biot ype	Analyzed cell lines		Analyzed human tissues									References					
						Leukocytes			Mammary			Adipose			Leuko.		Mam.		Tu	
				MCF7	mix	Non-cancer controls	BC patients	BRCA1 mut. carriers	Non-cancer controls	BC patients	BRCA1 mut. carriers	Non-cancer controls	BC patients	BRCA1 mut. carriers	Colombo (2014)	Romero (2015)	Colombo (2014)	Romero (2015)	Romero (2015)	Orban Olah (2003)
1Aq	**c.-25_-20del6**	**UTR**	**SDSΔ**	1185	7854	3632	5045	4210	2981	11439	4010	2508	6035	4955	**P**	**P**	**P**	**P**	**P**	**1**
1Aq, 2a	c.-25_-20del6, c.-19-59_-19-1 ins59	UTR	SDSΔ + SAS ▼	0	0	0	0	0	30	0	0	0	0	0	0	0	0	0	0	0
1Aq, Δ2	**c.-25_80del105**	**n.c.**	**SDSΔ + CΔ**	176	5032	40	101	150	46	191	28	59	82	205	**1**	**1**	**1**	**0**	**0**	**0**
1Aq, Δ2_3	c.-25_134del159	n.c.	SDSΔ + mCΔ	0	1575	0	202	13	14	46	0	0	0	0	1	1	1	0	0	0
1Aq, Δ2_5	c.-25_212del237	n.c.	SDSΔ + mCΔ	0	438	0	0	0	2	0	83	0	0	0	1	1	1	0	0	0
1Aq, Δ2_5, 6p	c.-25_282del307	n.c.	SDSΔ + mCΔ + SASΔ	0	350	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1Aq, Δ2_7	c.-25_441del466	n.c.	SDSΔ + mCΔ	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0
1Aq, Δ2_7, 8p	c.-25_444del469	n.c.	SDSΔ + mCΔ + SASΔ	0	0	0	34	0	0	0	0	0	0	0	0	0	0	0	0	0
1Aq, Δ2_10	c.-25_670del695	n.c.	SDSΔ + mCΔ	0	438	0	0	7	0	0	0	0	0	0	1	1	1	0	0	0
1Aq, Δ2_17	c.-25_5074del5099	n.c.	SDSΔ + mCΔ	88	1225	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1Aq, Δ2_19	c.-25_5193del5218	n.c.	SDSΔ + mCΔ	0	0	0	0	0	0	289	0	0	0	0	0	0	0	0	0	0
1aA	**c.-20+1_-20+89ins89**	**UTR**	**SDS ▼**	6824	12098	321	219	1211	1501	6557	5429	940	427	349	**1**	**1**	**1**	**0**	**P**	**0**
2p	**c.-19_-7del13**	**UTR**	**SASΔ**	88	963	140	118	232	77	214	28	76	238	164	**1**	**P**	**1**	**0**	**P**	**0**
Δ2	**c.-19_80del99**	**n.c.**	**CΔ**	176	3982	482	1063	1987	323	3305	1337	481	1308	462	**1**	**P**	**1**	**P**	**P**	**0**
Δ2_3	c.-19_134del153	n.c.	mCΔ	0	1575	0	0	0	36	364	0	5	0	82	1	P	1	0	P	0
Δ2_3, ▼4	c.-19_134del153 + c.135-4047_135-3932ins116	n.c.	mCΔ + Cq	110	0	0	0	0	0	0	0	5	0	51	1	1	0	0	0	0
Δ2_5	c.-19_212del231	n.c.	mCΔ	88	1203	20	0	0	73	35	0	0	0	0	1	1	1	0	0	0
Δ2_7, 8p	c.-19_444del463	n.c.	mCΔ + SASΔ	0	2253	10	169	0	0	0	0	0	0	0	0	0	0	0	0	0
Δ2_10	c.-19_670del689	n.c.	mCΔ	0	88	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1
Δ2_17	c.-19_5074del5093	n.c.	mCΔ	0	2625	0	0	0	0	0	0	11	0	0	0	0	0	0	0	0
Δ2_19	c.-19_5193del5212	n.c.	mCΔ	0	0	0	101	0	0	976	0	0	22	0	0	0	0	0	0	0
▼ 145 bp int 2	**c.81-3486_81-3342ins145**	**FS**	**C▼**	636	459	40	34	72	24	87	276	86	100	82	**0**	**0**	**0**	**0**	**0**	**0**
Δ3	**c.81_134del54**	**FS**	**CΔ**	987	21089	3120	10293	8436	4141	9336	9438	2135	5217	6750	**1**	**P**	**1**	**0**	**P**	**1**
Δ3, ▼4	c.81_134del54 + c.135-4047_135-3932ins116	FS	CΔ + C ▼	22	44	0	0	0	44	0	0	0	0	0	0	0	0	0	0	0
Δ3_5	c.81_212del132	IF	mCΔ	0	875	60	405	215	109	474	165	108	0	0	0	0	0	0	0	0
▼ 116bp int 3	c.134+3124_134+3239ins116	FS	C▼	461	1663	20	0	26	61	23	83	38	212	62	0	0	0	0	0	0
▼ 4	**c.135-4047_135-3932ins116**	**FS**	**C▼**	7460	2188	281	489	238	367	306	276	470	279	574	**1**	**1**	**1**	**0**	**P**	**0**
Δ5	**c.135_212del78**	**IF**	**C▼**	395	7351	8016	13043	8569	5524	7476	10169	4778	4013	8823	**P**	**1**	**P**	**P**	**P**	**1**
Δ5_6	c.135_301del167	FS	mCΔ	0	0	0	0	0	0	0	634	0	0	0	0	0	0	0	0	0
Δ5_6, 7p	c.135_307del173	FS	mCΔ + SASΔ	0	0	0	0	333	0	127	0	0	0	0	0	0	0	0	0	0
Δ5_7	c.135_441del307	FS	mCΔ	0	481	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Δ5_7, 8p	c.135_444del310	FS	mCΔ + SASΔ	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Δ5_9	c.135_593del459	IF	mCΔ	0	44	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5q	**c.191_212del22**	**FS**	**SDSΔ**	549	1269	3602	2818	2137	8023	4501	3624	5886	9339	4268	**P**	**P**	**P**	**P**	**P**	**1**
5q, Δ6	c.191_301del111	IF	SDSΔ + CΔ	0	0	0	0	0	0	0	0	22	0	0	0	0	0	0	0	1
Δ6	c.213_301del89	FS	CΔ	44	0	40	0	0	65	35	0	0	19	0	0	0	0	0	0	0
6q	**c.293_301del9**	**IF**	**SDSΔ**	0	44	431	371	695	202	116	28	124	7	62	**0**	**0**	**0**	**0**	**0**	**0**
8p	**c.442_444del3**	**IF**	**SASΔ**	197	7022	15079	10883	8566	7283	4287	4657	6356	5024	4011	**P**	**P**	**P**	**P**	**P**	**1**
Δ8	c.442_547del106	FS	CΔ	0	0	10	0	0	0	0	55	92	0	0	0	0	0	0	0	0
Δ8_9	c.442_593del152	FS	mCΔ	0	0	0	0	20	141	0	0	0	0	0	1	1	0	0	P	0
Δ8_10	c.442_670del229	FS	mCΔ	0	88	0	0	0	2	0	0	0	0	0	1	1	1	0	P	0
Δ8_16	c.442_4986del4545	IF	mCΔ	88	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
▼ 94 bp int 8	c.548-297_548-204ins94	FS	C▼	110	481	251	202	457	71	17	69	238	0	154	0	0	0	0	0	0
▼ 97 bp int 8	**c.548-300_548-204ins97**	**FS**	**C▼**	461	438	371	456	757	244	312	248	286	204	92	**0**	**0**	**0**	**0**	**0**	**0**
Δ9	**c.548_593del46**	**FS**	**CΔ**	1295	9888	3642	12334	17110	2295	7303	9301	2092	7042	8689	**P**	**1**	**P**	**0**	**P**	**1**
Δ9_10*	**c.548_670del123**	**IF**	**mCΔ**	4542	19558	562	1063	2261	998	2155	3362	940	1561	5109	**P**	**P**	**P**	**P**	**P**	**1**
Δ9_11	c.548_4096del3549	IF	mCΔ	1009	16079	100	270	454	2	0	413	22	0	923	P	1	P	0	0	1
Δ9_12	c.548_4185del3638	FS	mCΔ	0	0	10	0	0	0	0	165	16	0	0	1	1	1	0	0	0

10a	**c.594-21_594-1ins21**	**IF**	**SAS▼**	4827	2056	1575	5113	594	430	4957	1157	665	465	1036	**0**	**0**	**0**	**0**	**0**	**0**
Δ10	c.594_670del77	FS	CΔ	154	1684	30	34	0	12	0	0	27	7	0	1	1	1	0	P	0
Δ10_11	c.594_4096del3503	FS	mCΔ	0	4922	70	101	131	22	0	0	0	0	513	1	1	1	0	0	0
Δ10_12	c.594_4185del3592	FS	mCΔ	0	416	0	0	0	0	0	0	0	89	0	1	1	1	0	0	0
Δ11	**c.671_4096del3426**	**IF**	**CΔ**	439	5163	60	337	104	69	185	303	265	71	246	**1**	**1**	**1**	**0**	**0**	**1**
Δ11_12	c.671_4185del3515	FS	mCΔ	0	0	0	0	0	10	23	0	0	0	0	1	0	1	0	0	0
Δ11_12, 13p	c.671_4188del3518	FS	mCΔ + SASΔ	0	175	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0
11q	**c.788_4096del3309**	**IF**	**SDSΔ**	6012	16451	823	928	1965	2971	1491	5883	1048	2557	7879	**P**	**1**	**P**	**0**	**0**	**1**
11 Δ3094	c.788_3881del3094	FS	Intronization	241	219	0	0	39	6	0	0	0	0	0	0	0	0	0	0	0
11 Δ3110	c.788_3897del3110	FS	Intronization	2041	700	40	67	7	32	0	220	32	223	41	1	1	0	0	0	0
11 Δ3240	c.788_4027del3240	IF	Intronization	614	0	80	34	0	0	0	0	0	0	0	1	1	0	0	0	0
13p	**c.4186_4188del3**	**IF**	**SASΔ**	263	1378	3652	3881	9294	2684	560	3114	3297	847	6217	**P**	**1**	**P**	**0**	**P**	**0**
Δ13	**c.4186_4357del172**	**FS**	**CΔ**	44	1006	70	405	111	36	69	138	86	141	482	**1**	**1**	**0**	**0**	**P**	**0**
Δ13, 14p	c.4186_4360del175	FS	CΔ + SASΔ	44	219	20	0	0	6	12	0	0	0	0	1	1	0	0	0	0
Δ13_14	c.4186_4484del299	FS	mCΔ	0	2231	0	675	333	0	0	0	76	0	0	0	0	0	0	0	0
Δ13_15	c.4186_4675del490	FS	mCΔ	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0
▼ 13A	**c.4358-2785_4358-2720ins66**	**IF**	**C▼**	6495	24721	3963	1974	7267	873	1196	703	1254	5953	4309	**1**	**1**	**1**	**0**	**P**	**0**
▼ 13A, 14p	**c.4358-2785_4358-2720ins66 + c.4358_4360del3**	**IF**	**C▼ + SASΔ**	6604	40494	130	169	166	75	12	14	49	123	21	**1**	**1**	**1**	**0**	**0**	**0**
▼ 13A, Δ14	c.4358-2785_4358-2720ins66 + c.4358_4484del127	FS	C▼ + CΔ	0	0	0	0	819	115	0	0	0	0	0	0	0	0	0	0	0
14p	**c.4358_4360del3**	**IF**	**SASΔ**	4125	3238	542	1282	917	333	335	358	303	892	533	**P**	**P**	**P**	**P**	**P**	**0**
Δ14	**c.4358_4484del127**	**FS**	**CΔ**	417	1488	1304	877	1429	607	1231	1529	773	1442	882	**1**	**0**	**0**	**0**	**0**	**0**
Δ14_15	c.4358_4675del318	IF	mCΔ	241	1663	10	67	52	44	0	55	38	0	144	1	1	0	0	0	0
Δ14_17	c.4358_5074del717	IF	mCΔ	132	919	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1
Δ15	**c.4485_4675del191**	**FS**	**CΔ**	285	766	642	911	1723	220	254	468	168	286	1231	**1**	**1**	**0**	**0**	**P**	**0**
Δ15_16	c.4485_4986del502	FS	mCΔ	0	1619	30	0	274	0	0	0	0	0	0	0	0	0	0	0	0
Δ15_17	c.4485_5074del590	FS	mCΔ	176	372	130	945	1214	8	162	55	103	0	718	1	1	1	0	P	1
Δ15_19	c.4485_5193del709	FS	mCΔ	88	744	0	0	228	4	0	0	0	0	0	1	1	1	0	0	0
Δ15_23	c.4485_5467del983	FS	mCΔ	0	175	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16a	c.4986+1_4986+65ins65	FS	SDS▼	570	44	5568	0	6687	944	2461	3541	740	4292	0	0	0	0	0	0	0
Δ17	**c.4987_5074del88**	**FS**	**CΔ**	88	1466	2358	337	1951	1265	503	7854	1762	1988	6217	**1**	**1**	**0**	**0**	**P**	**0**
Δ17_18	c.4987_5152del166	FS	mCΔ	0	613	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Δ17_19	c.4987_5193del207	IF	mCΔ	88	1006	0	0	75	0	532	262	0	0	0	0	0	0	0	0	0
Δ17_20	c.4987_5277del291	IF	mCΔ	0	0	0	0	0	107	0	0	0	0	0	0	0	0	0	0	0
Δ18	c.5075_5152del78	IF	CΔ	0	219	60	0	0	44	0	0	0	0	0	1	1	0	0	0	0
Δ18_19	c.5075_5193del119	FS	mCΔ	0	0	10	67	0	4	81	28	11	7	0	0	0	0	0	0	0
Δ18_20	c.5075_5277del203	FS	mCΔ	0	0	30	0	0	4	0	0	0	0	0	0	0	0	0	0	0
▼ 143 bp int 19	c.5194-1231_5194-1089ins143	FS	C▼	132	0	20	0	0	381	324	138	573	0	431	0	0	0	0	0	0
▼ 146 bp int 19	c.5194-1234_5194-1089ins146	FS	C▼	373	197	80	101	653	40	0	289	32	19	421	0	0	0	0	0	0
Δ20	c.5194_5277del84	IF	CΔ	0	481	191	607	0	60	0	2618	0	0	0	1	1	1	0	0	0
Δ21	**c.5278_5332del55**	**FS**	**CΔ**	1711	8226	492	776	1765	389	376	1226	195	22	1057	**1**	**1**	**1**	**0**	**P**	**0**
Δ21_22	c.5278_5406del129	IF	mCΔ	88	438	0	0	0	0	549	0	0	0	0	1	1	1	0	P	0
▼ 129 bp int 21	**c.5332+873_5332+1001ins129**	**FS**	**C▼**	1711	263	592	641	1250	230	618	400	432	342	657	**0**	**0**	**0**	**0**	**0**	**0**
▼ 119 bp int 21	c.5333-706_5333-588ins119	FS	C▼	66	328	90	0	65	28	52	179	11	45	318	0	0	0	0	0	0
Δ22	**c.5333_5406del74**	**FS**	**CΔ**	4739	16451	1585	4590	3064	4405	11919	7372	3464	3872	5160	**1**	**P**	**1**	**0**	**P**	**0**
23a	c.5407-9_5407-1ins9	IF	SAS▼	44	66	10	0	39	35	168	55	68	74	0	0	0	0	0	0	0
Δ23	c.5407_5467del61	FS	CΔ	263	613	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0
Sum of identified variants in each sample (mPCR pool)				***54***	***71***	***56***	***46***	***51***	***64***	***49***	***48***	***50***	***40***	***41***						
Number of identified variants in analyzed cell lines / tissue type					***76***		***67***			***72***			***54***							

potentially result in a translation of BRCA1 protein isoforms altering its biological functions, as demonstrated for Δ14_15 and Δ17_19 previously (Sevcik et al., 2012; Sevcik et al., 2013). Nine out of 11 "ubiquitous" in-frame variants (Δ5; 8p; Δ9_10; Δ11; 11q; 13p; ▼13A; ▼13A, 14p; 14p) represented known ASVs while 6q and the highly expressed 10a variant were surprisingly not scored previously. The remaining four of the 25 in-frame variants (Δ3_5; Δ9_11; Δ14_15; 23a) were detected in most of the analyzed samples. Interestingly, Δ17_19 was detected in the cell lines and mammary tissue samples of *BRCA1*-mutation carriers and BC patients, but in no control sample.

Our approach enabled a direct quantification of 29 SDS/SAS variants (including 17 mixed biotypes). Twelve SDS/SAS-only biotype variants were identified. Most of them (10 out of 12) were identified as "ubiquitous", and only the 16a and 23a splicing variants were not present in some analyzed samples. Out of 17 more complex splicing mRNA BRCA1 isoforms (mixed biotypes), only 1Aq, Δ2 and ▼13A, 14p were "ubiquitous", while the other variants occurred rather rarely (Supplementary Table S6). Besides the 11q splicing variant, which lacks 3,309 bp from exon 11 and its identification was done with a specific forward primer, the other 28 relatively short indels were co-amplified stoichiometrically alongside the corresponding canonical splicing variant (Fig. 1H) and therefore we were able to quantify their relative

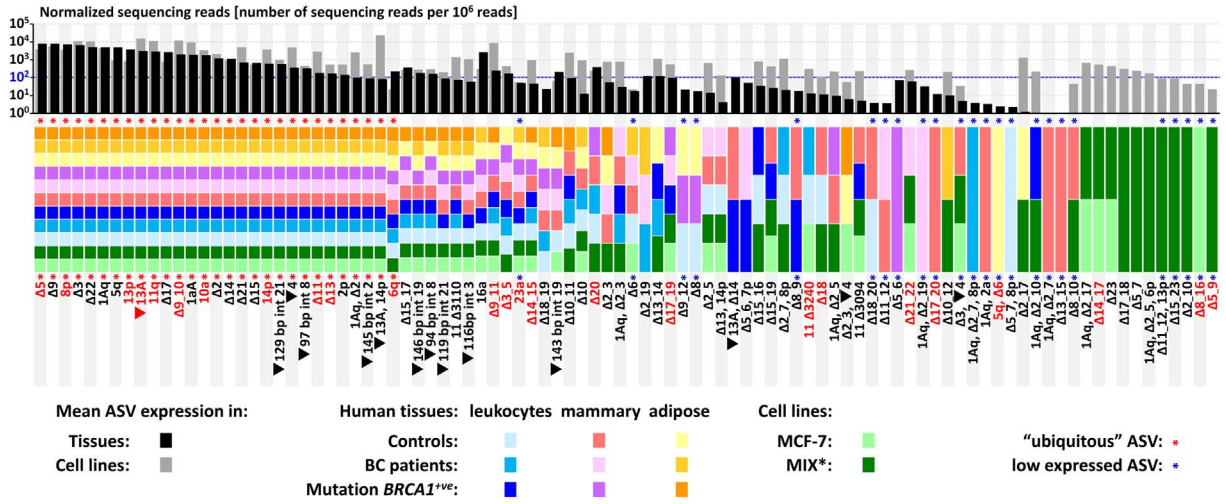


Fig. 2. Qualitative description of the presence of 94 identified BRCA1 ASVs (red letters indicate in-frame variants) in analyzed cDNA sample pools (colour bars). The grey-scale graph (upper part) shows the mean expression (in normalized reads per 10^6 reads) in human tissues (black) and cell line samples (grey). Red asterisks indicate ubiquitous variants (Table 1); blue asterisks indicate variants with low expression ($< 10^2$ normalized reads averaged in both tissue and cell samples).
*Mix of cDNAs from EM-G3, HeLa, and MDA-MB-231 cell lines. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

expression (compared with other existing variants in the particular exon-exon junction including wt transcript) directly from sequencing reads.

To test biological reproducibility, we compared the presence of BRCA1 ASVs in two independently analyzed sets of control samples, each consisting of leukocytes, mammary tissue and adipose tissue cDNA sample pools from eight individuals (Fig. 3A). Altogether, 74/94 BRCA1 ASVs were identified in at least one tissue control sample, while 20/94 were not present in any of them. Thirty-five variants (Fig. 3B) were consistently present (or absent – Δ17_20 in leukocytes and adipose tissue, 11Δ3240 in mammary and adipose tissue, and Δ20 in adipose tissue) in the analyzed tissue type biological replicate. These 35 variants included the majority of "ubiquitous" ASVs as they were detected with high mean coverage per variant and sample (371 and 1496 in absolute and normalized sequencing reads, respectively). The remaining 39 variants (shown in Fig. 3C), discordantly expressed in paired biological replicates, represented low expressed events with low mean coverage per variant and sample (11 and 41 in absolute and normalized sequencing reads, respectively). The differences in sequencing coverage between 35 consistently present and 39 discordantly expressed variants are shown in Fig. 3D.

4. Discussion

An accurate description of 'naturally occurring' ASVs is a prerequisite to understanding their biological significance. RNA-sequencing (RNA-seq) of human RNA samples revealed that 90% of multi-exon genes undergo alternative splicing (Wang et al., 2015). While RNA-seq represents a superior tool for qualitative and quantitative transcriptome analyses, including ASV identification (Byron et al., 2016), it is unsuitable for small-scale projects targeting a few or a single gene. Moreover, RNA-seq analyses of low expressed transcripts require the sequencing of up to 100 million mapped reads and sophisticated bioinformatics instruments (Wang et al., 2015; Conesa et al., 2016).

A pioneering systematic description of BRCA1 ASVs was made by Orban and Olah (2003) who reviewed 23 BRCA1 ASVs known in 2003. Recently, Colombo et al. (2014) identified 63 BRCA1 ASVs by an analysis of 38 blood-derived samples and one healthy breast tissue sample. Subsequently, Romero et al. (2015) revealed 54 BRCA1 ASVs in an analysis of 70 breast tumor samples, four breast samples from healthy individuals and 72 blood-derived samples. Two later studies described the characterization of BRCA1 ASVs by capillary electrophoresis, which required further cloning or sequencing of fragments containing splice junction events in order to identify the presented peaks. However, only the in silico imputation has been used to explain the peak pattern observed in capillary electrophoresis for a subset of events (Colombo et al., 2014).

Overall, 42 out of 94 BRCA1 splicing events described by our approach had not been identified in previous studies (Supplementary Table S5) which we used to compare the obtained results (Colombo et al., 2014; Romero et al., 2015; Orban & Olah, 2003).

The most common biotypes identified in our study (59/94; 63% variants) were cassette and multicassette ASVs (Supplementary Table S5). We found 27 cassette ASVs that included all 17 variants described in Colombo's study and 10 novel variants (eight intron exonizations, Δ6 and Δ8). Of 32 multicassette biotype ASVs ascertained in our study, 16 were described previously. We did not detect four multicassette ASVs (Δ14_18; Δ14_19; Δ21_23; Δ22_23) reported by Colombo's study as minor variants.

The second most frequent biotype variants were SDS/SAS (12/94; 13% variants). Besides nine described by Colombo et al. (2014), we found another four variants containing the exonizations of adjacent intronic regions ("ubiquitous" in-frame 6q and 10a, rare in-frame 23a, and a frameshift 16a) in all analyzed patient tissue types.

We found three large intronizations affecting exon 11, including two described by Colombo et al. (2014) previously, and the sparsely expressed frameshift variant 11Δ3094. We did not target terminal modifications involving the alternative exon 1B and IRIS in our analysis.

Furthermore, we recorded 20 mixed biotype variants including two "ubiquitous" (1Aq, Δ2 and ▼13A, 14p). Nine mixed biotype variants were described previously and 11 rare were novel. We did not find six variants detected in Colombo's study, including three variants with untested alternative exon 1B, and another three (1Aq, 2p; 1Aq, Δ2_3, ▼4; and Δ10_13p) previously described as minor.

The most complex SAS/SDS events affected the non-coding 5′ untranslated region (at the exon 1A-2 junction). The 1Aa variant containing an insertion of 89 nucleotides prevailed in cell line samples. The dominant variant in all analyzed tissue samples was wt exon 1A accompanied by the shortened variant exon 1Aq (in approximately one-third of all mapped reads). The expression of three other SAS variants 8p; 14p; 13p (lacking the CAG nucleotides at the 5′-end of an exon) was ~10% of all sequencing reads in most of the analyzed samples. These variants rank among the NAGNAG tandem acceptors, a common kind of ASVs resulting in single amino acid exclusion (Sinha et al., 2009). The

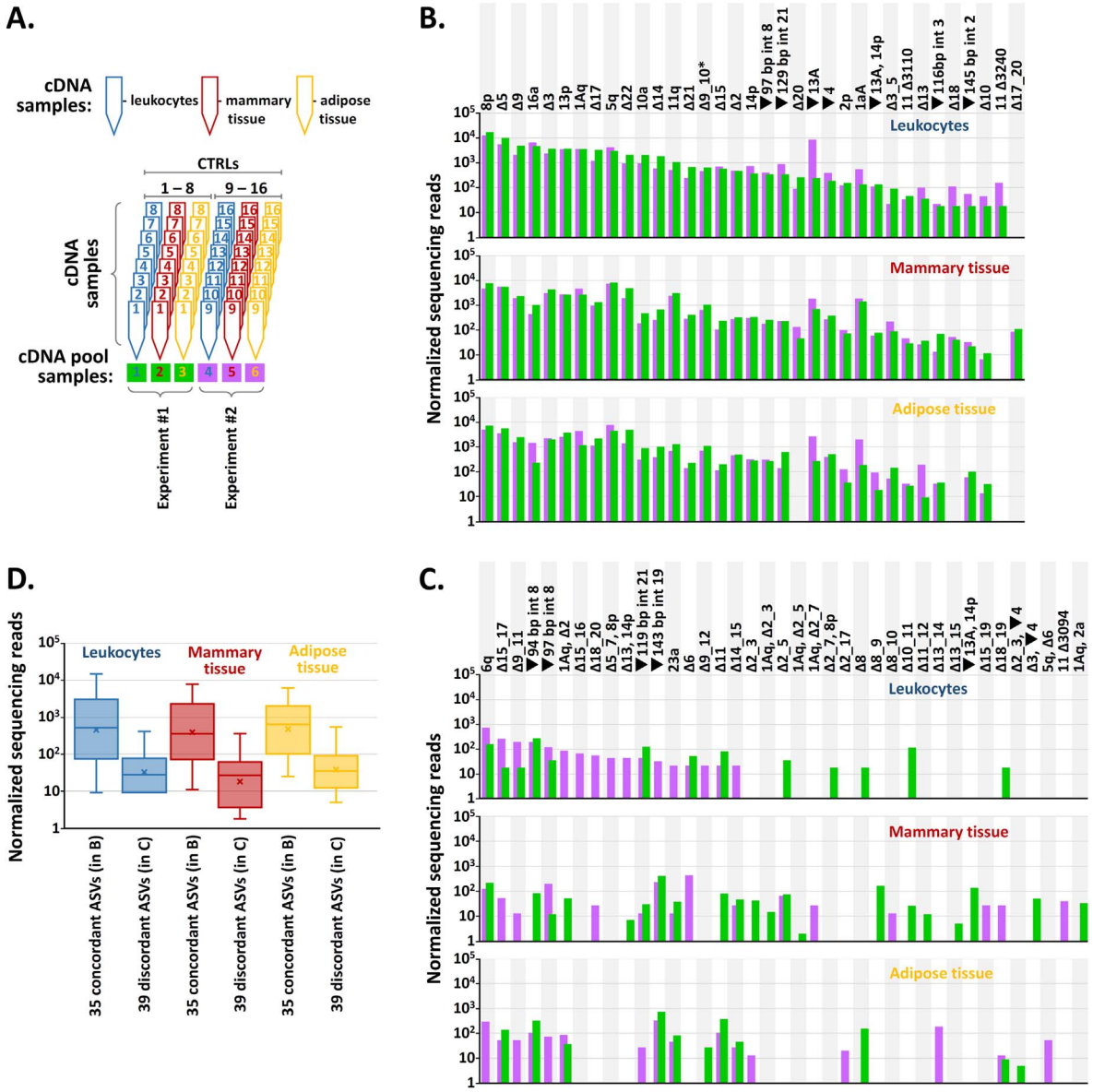


Fig. 3. The reproducibility test of the method in biological replicates involved an analysis of independent cDNAs from three types of tissue obtained from 16 control individuals. Panel A (adjusted from Fig. 1C) shows the arrangement of experiments #1 and #2, each consisting of NGS analyses from leukocytes, mammary tissue, and adipose tissue proceeded and sequenced independently by a pipeline described in the Method section. Graphs B and C compare the numbers of normalized sequencing reads (in log scale) for BRCA1 ASVs (listed in Supplementary Table 4) expressed in leukocytes, mammary and adipose tissues in two independent sets (Experiment #1 and #2) of control samples obtained from 16 individuals. The samples from control individuals 1–8 were analyzed in Experiment #1 (green bars), while the samples from control individuals 9–16 were analyzed in Experiment #2 (violet bars). Panel B shows the expression of 35 fully reproducible ASVs that gave concordant results in all analyzed biological duplicates. Panel C shows the expression of 39 non-fully reproducible ASVs that gave discordant results in at least one biological duplicate. The expressions of 35 fully reproducible ASVs were substantially higher than the expression of 39 non-fully reproducible ASVs as shown in panel D. The box plot charts show the values of sequencing normalized coverage (in log10 scale) for 35 fully reproducible ASVs (shown in B) and 39 non-fully reproducible ASVs (shown in C) in the analyzed tissues. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

other variants were minor or rare, with the exception of five ASVs (10a; 16a; 1Aq, Δ2; Δ13, 14p; and ▼13A, 14p) expressed with a higher proportion in cell line samples.

The presented approach showed satisfactory reproducibility documented at the level of mPCR amplification (Fig. 1E and G) and also in the analysis of biological replicates from two sets of tissue samples from control individuals (Fig. 3). We suppose that discrepancies in the detection of individual ASVs in biological replicates resulted rather from differentially expressed BRCA1 ASVs in the individual RNA samples (mixed in the mPCR pools) than from analytical variability, because reproducibility was strongly positively correlated with the level of variant expression (i.e. coverage; Fig. 3D). Our analysis, performed in native tissues and cells (with unmodified nonsense-mediated decay pathway), revealed 48 frame-shift variants. However, their ratio to in-frame variants was strongly reduced in the subset of "ubiquitous" variants (48/25 in the entire set of BRCA1 ASVs and 13/11 in "ubiquitous" variants). It has to be noted that BRCA1 mRNA is expressed at low levels, a few tenths of copies per cell in MCF7 cells (Lee et al., 2014). Many of newly identified ASVs were represented by a low number of reads (Fig. 2), indicating that they were probably expressed in a few copies per cell or present only in a subset of cDNA samples in the analyzed cDNA pool. We suppose that at least some of them may represent stochastic noise determining the number of alternative isoforms and their abundance (Melamud & Moult, 2009).

Deletions in isoform 1 (NM_007294.3) are the most frequently described BRCA1 ASVs. The longest well described ASV intron exonization is ▼4 (116 bp from intron 3). The predicted length of PCR amplicon covering ▼4 was 263 bp in our analysis, while the shortest

predicted amplicons (covering variants Δ8_17; 11q, Δ12_17; Δ13_17; Δ17; Δ19_21) were 60 bp long. Therefore, the setup of mPCR and the size-selection protocol were targeted to enrich amplicons with mean fragment length of 100–150 bp in order to disclose majority of putative ASVs. We were able to identify splicing events covered by PCR fragments ranging between 60 and 287 bp. One of the shortest identified amplicons was "ubiquitous" variant Δ17, while the variant 11Δ3094 was characterized from PCR product of 287 bp (the above mentioned variant ▼4, identified from 263 bp amplicon, occurred as "ubiquitous"). These findings indicate the range of amplicon lengths (60–263 bp) which can be analyzed under defined conditions.

In conclusion, the mPCR/NGS approach enables direct identification of all biotype classes of splicing events, including mixed biotypes containing exonizations of flanking intronic sequences. The analysis of BRCA1 mRNA revealed the broadest spectrum of its splicing variants, including their distribution in the analyzed human tissue and cell line samples. Similar to most other methods (including recent RNA-seq analyses), the analysis is not able to identify possible combinations of splicing events affecting both 5′ and 3′ portions of the large BRCA1 transcripts. We are also aware that our approach could miss large deep intronic exonizations (substantially exceeding the targeted PCR amplification and/or range of size selection). We would like to emphasize that the described method can be easily adopted for an analysis of any gene of interest in order to identify its ASVs, not only in human samples. Additionally, we suppose that our approach may represent an interesting option for the functional classification of VUS introducing aberrant splicing with modified protocol using individual (instead of pooled) cDNA template for mPCR step (limited to the region of interest).

Funding

This work was supported by the Grant Agency of the Czech Republic [grant number P301-12-1850], Charles University projects [PROGRES Q28/LF1, and SVV2017/260367], and a Ministry of Education, Youth and Sports project [The National Center for Medical Genomics LM2015091].

Disclosure statement

The authors have no conflicts of interest to disclose.

Acknowledgements

We thank The National Center for Medical Genomics for their instrumental and technical support with NGS sequencing and data analyses.

Authors' contributions

Study design (ZK, PK), experimental procedures (JH, FL, HH, KH, MJ, JSo), bioinformatics (PZ, VS), sample collection and characterization (OM, DP, MV, JSe), data analysis (JH, PK), manuscript preparation (PK, ZK, JH, JSe), manuscript approval (all authors).

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.gene.2017.09.025.

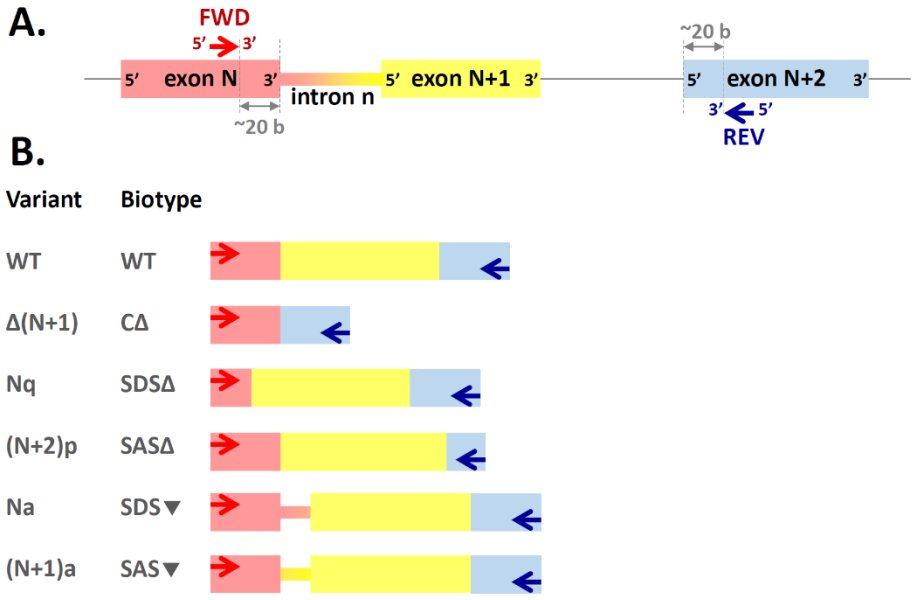
References

Baralle, D., Buratti, E., 2017. RNA splicing in human disease and in the clinic. Clin. Sci. (Lond.) 131, 355–368.
Bentley, D.L., 2014. Coupling mRNA processing with transcription in time and space. Nat. Rev. Genet. 15, 163–175.
Brozova, M., Kleibl, Z., Netikova, I., Sevcik, J., Scholzova, E., Brezinova, J., Chaloupkova, A., Vesely, P., Dundr, P., Zadinova, M., et al., 2007. Establishment, growth and in vivo differentiation of a new clonal human cell line, EM-G3, derived from breast cancer progenitors. Breast Cancer Res. Treat. 103, 247–257.
Bustin, S.A., Benes, V., Garson, J.A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M.W., Shipley, G.L., et al., 2009. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clin. Chem. 55, 611–622.
Byron, S.A., Van Keuren-Jensen, K.R., Engelthaler, D.M., Carpten, J.D., Craig, D.W., 2016. Translating RNA sequencing into clinical diagnostics: opportunities and challenges. Nat. Rev. Genet. 17, 257–271.
Caminsky, N., Mucaki, E., Rogan, P., 2015. Interpretation of mRNA splicing mutations in genetic disease: review of the literature and guidelines for information-theoretical analysis. F1000Res 18, 282.
Cheon, J.Y., Mozersky, J., Cook-Deegan, R., 2014. Variants of uncertain significance in BRCA: a harbinger of ethical and policy issues to come? Genome Med. 6, 121.
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.
Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., Cervera, A., McPherson, A., Szczesniak, M.W., Gaffney, D.J., Elo, L.L., Zhang, X., et al., 2016. A survey of best practices for RNA-seq data analysis. Genome Biol. 17, 13.
Kleibl, Z., Kristensen, V.N., 2016. Women at high risk of breast cancer: molecular characteristics, clinical presentation and management. Breast 28, 136–144.
Kleiblova, P., Dostalova, I., Bartlova, M., Lacinova, Z., Ticha, I., Krejci, V., Springer, D., Kleibl, Z., Haluzik, M., 2010. Expression of adipokines and estrogen receptors in adipose tissue and placenta of patients with gestational diabetes mellitus. Mol. Cell. Endocrinol. 314, 150–156.
Lee, K., Cui, Y., Lee, L.P., Irudayaraj, J., 2014. Quantitative imaging of single mRNA splice variants in living cells. Nat. Nanotechnol. 9, 474–480.
Melamud, E., Moult, J., 2009. Stochastic noise in splicing machinery. Nucleic Acids Res. 37, 4873–4886.
Orban, T.I., Olah, E., 2003. Emerging roles of BRCA1 alternative splicing. Mol. Pathol. 56, 191–197.
Rahman, N., 2014. Realizing the promise of cancer predisposition genes. Nature 505, 302–308.
Romero, A., Garcia-Garcia, F., Lopez-Perolio, I., Ruiz de Garibay, G., Garcia-Saenz, J.A., Garre, P., Ayllon, P., Benito, E., Dopazo, J., Diaz-Rubio, E., et al., 2015. BRCA1 alternative splicing landscape in breast tissue samples. BMC Cancer 15, 219.
Sevcik, J., Falk, M., Kleiblova, P., Lhota, F., Stefancikova, L., Janatova, M., Weiterova, L., Lukasova, E., Kozubek, S., Pohlreich, P., et al., 2012. The BRCA1 alternative splicing variant Delta14-15 with an in-frame deletion of part of the regulatory serine-containing domain (SCD) impairs the DNA repair capacity in MCF-7 cells. Cell. Signal. 24, 1023–1030.
Sevcik, J., Falk, M., Macurek, L., Kleiblova, P., Lhota, F., Hojny, J., Stefancikova, L., Janatova, M., Bartek, J., Stribrna, J., et al., 2013. Expression of human BRCA1Delta17-19 alternative splicing variant with a truncated BRCT domain in MCF-7 cells results in impaired assembly of DNA repair complexes and aberrant DNA damage response. Cell. Signal. 25, 1186–1193.
Sinha, R., Nikolajewa, S., Szafranski, K., Hiller, M., Jahn, N., Huse, K., Platzer, M., Backofen, R., 2009. Accurate prediction of NAGNAG alternative splicing. Nucleic Acids Res. 37, 3569–3579.
Sloan, C.A., Chan, E.T., Davidson, J.M., Malladi, V.S., Strattan, J.S., Hitz, B.C., Gabdank, I., Narayanan, A.K., Ho, M., Lee, B.T., et al., 2016. ENCODE data at the ENCODE portal. Nucleic Acids Res. 44, D726–732.
Soukarieh, O., Gaildrat, P., Hamieh, M., Drouet, A., Baert-Desurmont, S., Frebourg, T., Tosi, M., Martins, A., 2016. Exonic splicing mutations are more prevalent than currently estimated and can be predicted by using in silico tools. PLoS Genet. 12, e1005756.
Tavtigian, S.V., Chenevix-Trench, G., 2014. Growing recognition of the role for rare missense substitutions in breast cancer susceptibility. Biomark. Med 8, 589–603.
Vondruskova, E., Malik, R., Sevcik, J., Kleiblova, P., Kleibl, Z., 2008. Long-term BRCA1 down-regulation by small hairpin RNAs targeting the 3′ untranslated region. Neoplasma 55, 130–137.
Wang, J., Ye, Z., Huang Tim, H.M., Shi, H., Jin, V., 2015. A survey of computational methods in transcriptome-wide alternative splicing analysis. Biomol. Concepts 6, 59–66.
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.

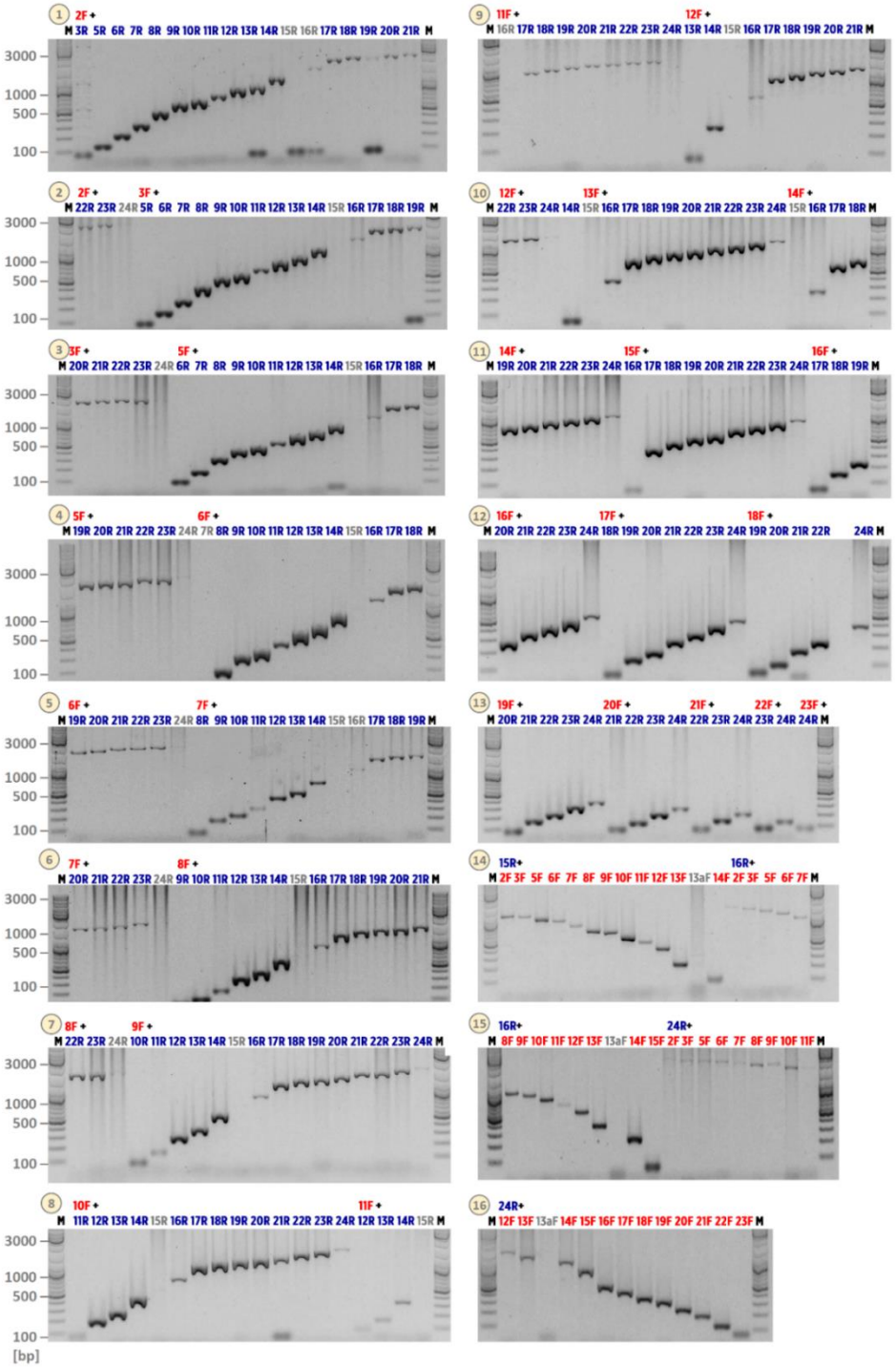
Supplementary Figure 1. Schematic depiction of forward (FWD in exon N) and nearest reverse (REV in exon N+2) primer localization for mPCR analysis (A). The enrichment of amplicons from alternatively spliced mRNA (cDNA) templates (in this case Δ(N+1) in B) is achieved by an exclusion of the reverse primer in the exon N+1 (an exon immediately following exon with forward primer). The presence of longer amplicon (WT in B) generated from wt templates (products of canonical splicing) is limited by a size selection in the subsequent step (not shown).

The 3' end of each primer was designed to localize approximately 20-30 b (~20 b in A) before 5' end of targeted exon (considering the DNA sequence) i.e. before nearest exon-exon junction (in mRNA). This distance reflects possible presence of splice donor/acceptor shifts exclusions (SDSΔ/SASΔ) at the exon-exon junction. The length from exon-exon junction can be modified (extended) according to the optimal length of expected mPCR amplicons from alternatively spliced mRNA (cDNA) templates (e.g. Δ(N+1)) as required for production of sequencing libraries in NGS analysis. The above mentioned parameters (~20 b in A) were used for NGS analysis with MiSeq Reagent Kit v3 (150 cycles 2x75 bp). However, the gap between 3' end of the primer and the exon-exon junction should be extended to 50-70 b with the use of sequencing kits reading longer amplicons. Ideally, all primers for mPCR amplification should have similar annealing temperature.

These considerations in primer design enable to amplify various splicing biotypes (B) including cassette exon N+1 exclusion (CΔ), splice donor/acceptor shifts exclusion (SDSΔ/SASΔ) or inclusions (SDS▼/SAS▼) and mixed biotypes (not shown).



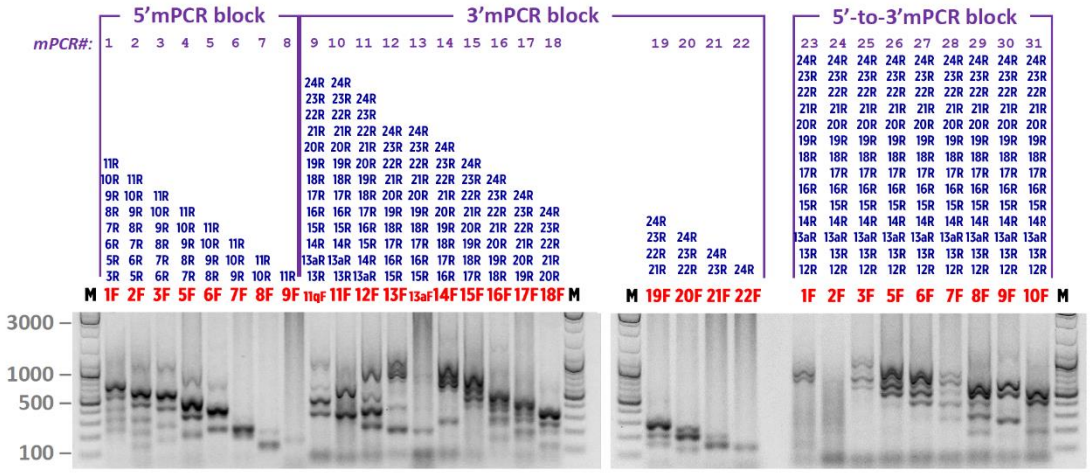
Supplementary Figure 2. Agarose gel electrophoresis of individual PCR amplicons optimized prior mPCR setup (panels 1-13). The reactions contained a single forward primer (red letters) and a single reverse primer (blue letters) as indicated. Forward primers yielding an inferior PCR amplification (gray letters) were redesigned to improve reaction efficiency and re-evaluated by control PCR reactions (panels 14-16). As the template for PCR optimization, we used the plasmid construct pcDNA3.1-HA containing BRCA1 coding sequence (described in Sevcik et al.). Therefore, the amplicons targeting alternative exon 13 (13a in panels 14-16) not present in the plasmid template were not detected.



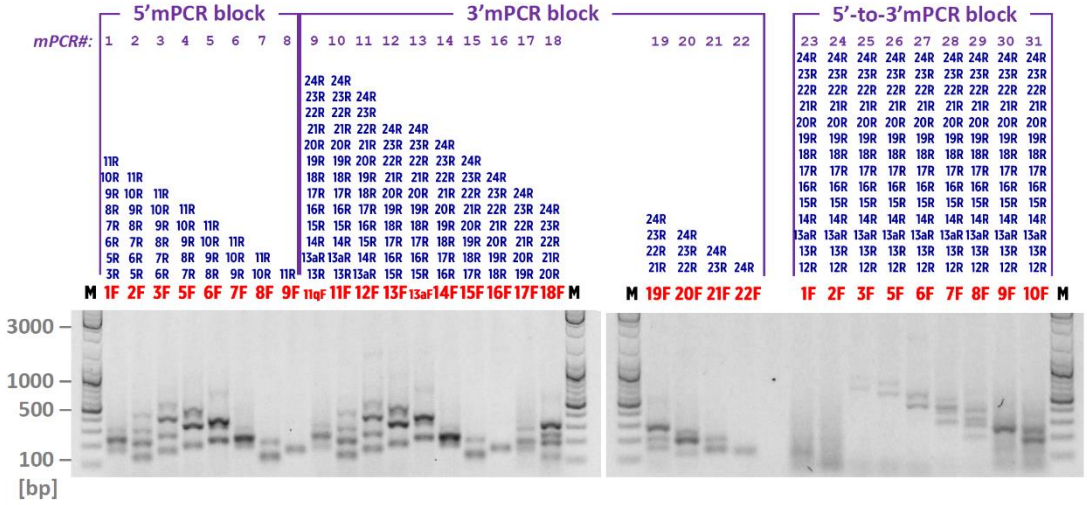
Sevcik J et al. Expression of human BRCA1Delta17-19 alternative splicing variant with a truncated BRCT domain in MCF-7 cells results in impaired assembly of DNA repair complexes and aberrant DNA damage response, Cell. Signal. 25 (2013) 1186-1193.

Supplementary Figure 3. Setup and optimization of mPCR reactions. The 31 mPCR reactions (mPCR#) were individualy amplified. In each reaction, a single forward primer (in red letters) was mixed with 1-14 reverse primers (blue letters) . The mPCR setup considered the presence of large exon 11 encompassing over 60% of canonical *BRCA1* coding sequence. Therefore, mPCRs were organized in three mPCR 'blocks' (5' mPCR block, 5'-to-3' mPCR block, and 3' mPCR block). The initial setup **(A)** with elongation time in mPCR reaction of 120 s was performed to control amplified mPCR amplicons from cDNA pool (consisting of equimolarly mixed cDNAs of EM-G3, HeLa, and MDA-MB-231 cells). This cDNA pool we used because of the expected presence of multiple BRCA1 mRNA alternative splicing variants. The routine mPCR amplification performed for sequencing library preparation **(B)** had reduced elongation time for mPCR reaction to 15 s in order to limit presence of udesired longer amplicons containing PCR products of BRCA1 canonical splicing. These amplicons were generated especially in reactions with multiple reverse primers (as could be seen e.g. for amplicons in 5'-to-3' block).

A. PCR elongation time: 120 s



B. PCR elongation time: 15 s



A note: the picture showing agarose gel electrophoresis in panel B is identical to that present in Figure 1D.

Příloha 5 – Plný text manuskriptu: Sevcik J, Falk M, Macurek L, Kleiblova P, Lhota F, Hojny J, Stefancikova L, Janatova M, Bartek J, Stribrna J, Hodny Z, Jezkova L, Pohlreich P, Kleibl Z. **Expression of human BRCA1Δ17-19 alternative splicing variant with a truncated BRCT domain in MCF-7 cells results in impaired assembly of DNA repair complexes and aberrant DNA damage response.** Cell Signal. 2013 May;25(5):1186-93. doi:10.1016/j.cellsig.2013.02.008

Cellular Signalling 25 (2013) 1186–1193

Contents lists available at SciVerse ScienceDirect

Cellular Signalling

journal homepage: www.elsevier.com/locate/cellsig

Expression of human BRCA1Δ17–19 alternative splicing variant with a truncated BRCT domain in MCF-7 cells results in impaired assembly of DNA repair complexes and aberrant DNA damage response

Jan Sevcik [a,e,*], Martin Falk [b], Libor Macurek [c,d], Petra Kleiblova [a], Filip Lhota [a], Jan Hojny [a], Lenka Stefancikova [b], Marketa Janatova [a], Jiri Bartek [c,f,g], Jana Stribrna [a], Zdenek Hodny [c], Lucie Jezkova [b,h], Petr Pohlreich [a], Zdenek Kleibl [a,**]

[a] Institute of Biochemistry and Experimental Oncology, First Faculty of Medicine, Charles University in Prague, U Nemocnice 5, CZ-128 53 Prague 2, Czech Republic
[b] Institute of Biophysics, The Academy of Sciences of the Czech Republic, Kralovopolska 135, CZ-61265 Brno, Czech Republic
[c] Laboratory of Genome Integrity, Institute of Molecular Genetics, v.v.i., The Academy of Sciences of the Czech Republic, Videnska 1083, CZ-14220 Prague, Czech Republic
[d] Laboratory of Cancer Cell Biology, Institute of Molecular Genetics, v.v.i., The Academy of Sciences of the Czech Republic, Videnska 1083, CZ-14220 Prague, Czech Republic
[e] Prague Burn Centre, Charles University, Third Faculty of Medicine and Teaching Hospital Kralovske Vinohrady, Srobarova 50, CZ-10034 Prague 10, Czech Republic
[f] Institute of Molecular and Translational Medicine, Faculty of Medicine and Dentistry, Palacky University Olomouc, Hnevotinska 5, CZ-779 00, Czech Republic
[g] Danish Cancer Society Research Center, Strandboulevarden 49, DK-2100 Copenhagen, Denmark
[h] Institute of Chemical Technology Prague, Faculty of Food and Biochemical Technology, Technicka 5, CZ-166 28 Prague 6, Czech Republic

ARTICLE INFO

Article history:
Received 14 January 2013
Accepted 8 February 2013
Available online 14 February 2013

Keywords:
BRCA1
Alternative splicing
Breast cancer
DNA damage
Homologous recombination (HR)
Non-homologous end joining (NHEJ)
BRCA1 C-terminal (BRCT) domain
BRCA1 containing complexes
Ionizing radiation-induced foci (IRIF)
Abraxas
CtIP
BARD1

ABSTRACT

Alternative pre-mRNA splicing is a fundamental post-transcriptional regulatory mechanism. Cancer-specific misregulation of the splicing process may lead to formation of irregular alternative splicing variants (ASVs) with a potentially negative impact on cellular homeostasis. Alternative splicing of BRCA1 pre-mRNA can give rise to BRCA1 protein isoforms that possess dramatically altered biological activities compared with full-length wild-type BRCA1. During the screening of high-risk breast cancer (BC) families we ascertained numerous BRCA1 ASVs, however, their clinical significance for BC development is largely unknown. In this study, we examined the influence of the BRCA1Δ17–19 ASV, which lacks a portion of the BRCT domain, on DNA repair capacity using human MCF-7 BC cell clones with stably modified BRCA1 expression. Our results show that overexpression of BRCA1Δ17–19 impairs homologous recombination repair (sensitizes cells to mitomycin C), delays repair of ionizing radiation-induced DNA damage and dynamics of the ionizing radiation-induced foci (IRIF) formation, and undermines also the non-homologous end joining repair (NHEJ) activity. Mechanistically, BRCA1Δ17–19 cannot interact with the partner proteins Abraxas and CtIP, thus preventing interactions known to be critical for processing of DNA lesions. We propose that the observed inability of BRCA1Δ17–19 to functionally replace wtBRCA1 in repair of DNA double-strand breaks (DDSB) reflects impaired capacity to form the BRCA1-A and -C repair complexes. Our findings indicate that expression of BRCA1Δ17–19 may negatively influence genome stability by reducing the DDSB repair velocity, thereby contributing to enhanced probability of cancer development in the affected families.

1. Introduction

BRCA1 (MIM# 113705) is a multifunctional protein and established tumor suppressor participating in the regulation of several distinct cellular processes. Its main activity involves mediating the DNA damage response [1,2]. BRCA1 participates in non-homologous end joining (NHEJ) and homologous recombination (HR) of DNA double-strand breaks (DDSB), where it acts as a protein–protein interaction platform [3,4]. Specific differential activities of BRCA1 within the DDSB repair process are determined by assembly of discrete multiprotein complexes, through association of various interacting proteins with the BRCA1 C-terminal BRCT domain and the N-terminal RING finger motif [5]. While the participation of BRCA1 in NHEJ is not mechanistically

Abbreviations: AS, alternative splicing; ASV, alternative splicing variant; B2M, β-2-microglobulin; BARD1, BRCA1-associated RING domain 1; BC, breast cancer; BRCA1/2, breast cancer-associated gene 1/2; BRCT, BRCA1 C-terminal domain; BRCC36, BRCA1/BRCA2-containing complex subunit 36; BRIP1, BRCA1-interacting protein 1; CtIP, C-terminal interaction protein; DDR, DNA damage response; DDSB, DNA double strand break; HR, homologous recombination; IRIF, ionizing radiation-induced foci; NHEJ, non-homologous end joining; PBGD, porphobilinogen deaminase; PE, plating efficiency; PI, post-irradiation; RING, Really interesting new gene (domain); RPA, replication protein A; RTCA, real-time cell analyzer; SCD, serine-containing domain; SF, surviving fraction; shRNA, short-hairpin RNA; ssDNA, single-stranded DNA; 53BP1, p53 binding protein 1.

* Correspondence to: J. Sevcik, Institute of Biochemistry and Experimental Oncology, First Faculty of Medicine, Charles University in Prague, U Nemocnice 5, CZ-128 53, Prague 2, Czech Republic. Tel.: +420 22496 5745; fax: +420 22496 5732.
** Corresponding author. Tel.: +420 22496 5745; fax: +420 22496 5732.
E-mail addresses: jsevc@lf1.cuni.cz (J. Sevcik), zdekleje@lf1.cuni.cz (Z. Kleibl).

0898-6568/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.cellsig.2013.02.008

understood, it is known that at least three distinct BRCA1-containing complexes (BRCA1 A, BRCA1 B and BRCA1 C, respectively) orchestrate HR. The core of the BRCA1 A-complex consists of the Abraxas (CCDC98) adaptor protein bound to BRCA1's BRCT domain. This complex possesses both ubiquitin ligase (provided by BRCA1/BARD1) and deubiquitinase (provided by BRCC36) activities [6]. It has been proposed that the ubiquitination of proteins acting in DDSB repair sustains a subtle balance in the rate of DNA resection at the DDSB site [7]. The BRCA1 B-complex is formed by binding BRIP1 (alias BACH1) helicase to BRCA1 and it is important for the extension of ssDNA regions and RPA loading during the resolution of stalled replication forks [8]. The BRCA1 C-complex contains CtIP exonuclease linked with BRCA1, and it represents the main enzyme to carry out 5′-to-3′ resection of free DNA ends at the DDSB site [9].

The step-wise formation and regulation of the BRCA1 complexes are not fully known yet, but these processes likely determine the precise timing and nature of BRCA1's activities in the DNA damage response [10]. The assembly of BRCA1 complexes can be limited by the spatio-temporal availability and binding capacity of BRCA1 and its binding partners. The nuclear targeting of BRCA1 itself is regulated by its nuclear localization and export signals, while its binding capacity is regulated by phosphorylation of the BRCA1 protein, particularly on its serine-containing domain (SCD) [11]. Changes in the BRCA1 protein structure especially in conservative domains broadly impair BRCA1's binding capacity and hence its ability to form functionally active complexes. Alterations in exons coding for conservative BRCA1 motifs frequently represent pathogenic mutations predisposing to breast, ovarian and other BRCA1-associated cancers in their carriers [12]. The structure of the BRCA1 protein could also be affected by altered pre-mRNA splicing (AS). Besides BRCA1 aberrant splicing (arising from alterations in conservative *cis*-regulatory splicing sequences), a number of BRCA1 alternative splicing variants (ASVs) have been reported [13,14]. While the functional and clinical significance of BRCA1 ASVs is mostly unknown, some variants show an obvious alteration in BRCA1 activities [15]. In general, pre-mRNA AS is considered as a means to increase genetic diversity by formation of protein isoforms with changed biological activities on the post-transcriptional level. Despite growing evidence of AS relevance for malignant transformation, there is a lack of knowledge about cancer-specific (de)regulation of AS and little mechanistic insight into malfunction of the various cancer-specific ASVs [16,17].

This work follows our previous study focused on the functional *in vitro* characterization of the BRCA1Δ14–15 ASV lacking a part of phosphorylation-targeted SCD. Here we report the functional characterization of the BRCA1Δ17–19 ASV, identified during the screening of high-risk breast cancer (BC) patients in the Czech Republic [18]. In particular, we describe the impact of the BRCA1Δ17–19 ASV on γ-radiation-induced DDSB repair complexes and repair capacity, using human MCF-7 BC cells as a model.

2. Material and methods

2.1. Construction of expression vectors

A pcDNA3.1-HA-wtBRCA1 construct was generously provided by Dr. Paul D. Harkin [19]. A pcDNA3.1-HA-BRCA1Δ17–19 expression construct (with in-frame deletion of exons 17–19 that code for a substantial portion of the first BRCT domain) was prepared by a PCR splicing approach using pcDNA3.1-HA-wtBRCA1 as a template, as described previously [15].

A pSUPER.retro.puro (Oligoengine) vector system was used for a short hairpin RNA (shRNA)-mediated downregulation of an endogenous wtBRCA1 expression. Particular pSUPER.retro.puro shRNA plasmids were constructed as described previously [20]. To assess only the biological activity of the BRCA1Δ17–19 ASV, the interfering RNAs were targeted to the BRCA1 exon 17 (shRNA 5196: 5′-GATCCCC**GTACAAGTTTGCCAGAAAA**TTCAAGAGATTTTCTGGCAAACTTGTACTTTTTGGAAC-3′) and 18 (shRNA 5331: 5′-GATCCCC**GAAAATGGGTAGTTAGCTA**TTCAAGAGATAGCTAACTACCCATTTTCTTTTTGGAAC-3′), respectively. Critical experiments were performed with both shRNAs to rule out an off-target effect of the used shRNA (Supplementary Fig. 1).

2.2. Cell culture and transfection

The MCF-7 cell line was used for functional analysis of the BRCA1Δ17–19 ASV, and the 293T cell line was used for an analysis of BRCA1 complexes with the BRCA1Δ17–19 ASV. Both cell lines were grown in DMEM (Gibco) medium supplemented with 10% fetal calf serum (Gibco), at 37 °C in a humidified atmosphere of 5% CO_2. The MCF-7 clones with stably modified expression of BRCA1 were selected as described previously [15,20]. Transient transfectants of 293T cells were prepared by polyethylenimine transfection as described [21].

2.3. Irradiation of cells

Non-synchronized cells grown on Petri dishes were irradiated with a single dose of 1.5 Gy of ionizing radiation using ^{60}Co (0.58 Gy/min) as described previously [15].

2.4. Comet assay

The MCF-7 clones grown in triplicates under standard conditions were γ-irradiated as described above. Non-irradiated cells (0 min) and cells at the particular times after irradiation (15, 30, 60, and 120 min) were assayed according to the manufacturer's instructions using a neutral CometAssay (Trevigen). The results were analyzed using the CometScore v1.5 (TriTec) software.

2.5. Mitomycin C sensitivity assay

An xCELLigence real-time cell analyzer (RTCA; Roche) was used for testing the sensitivity of clones to mitomycin C, as described previously [15]. The RTCA software ver. 1.2 (Roche) was used to analyze cell proliferation and dose response curves.

2.6. Immunofluorescence and high-resolution confocal microscopy

The MCF-7 clones grown under standard conditions were treated with γ-radiation as described above. Non-irradiated cells (0 min) and cells at the particular times (5, 30, 60, 120, 240 and 1440 min) after irradiation were fixed (4% paraformaldehyde, 10 min RT), permeabilized (0.2% Triton X100/PBS for 14 min RT) and immunoassayed for γH2AX/53BP1 colocalization as described previously, using the following antibodies: anti-phospho-H2AX (Upstate Biotechnology), rabbit anti-53BP1 (Cell Signaling), FITC-conjugated donkey anti-mouse and Cy3-conjugated donkey anti-rabbit antibodies (Jackson Laboratory), and TOPRO-3 (Molecular Probes) [15,22–24]. These IRIF markers – a phosphorylated form of the H2AX histone (γH2AX) and the DDSB mediator protein 53BP1 – co-localize at the site of DNA damage [11]. While γH2AX foci formation is generally accepted as a sufficient quantitative marker of DDSBs [25], the used double-labeling with 53BP1 enables a more precise quantification of DDSBs. The exposure time and dynamic range of the camera in the red, green and blue channels were adjusted to the same values for all slides to obtain quantitatively comparable images. Forty serial optical sections were captured at 0.2–0.3 μm steps along the z-axis, using the automated DM RXA confocal fluorescence microscope (Leica).

2.7. In vitro NHEJ assay

The overall capacity and precision of NHEJ was determined by a luciferase vector-based *in vitro* assay, as described previously [15]. Briefly, the DDSB was simulated by an enzymatic cleavage of a reporter pGL (Promega) vector at two different sites prior to transfection. Expression

of active luciferase from an EcoRI-linearized pGL vector in the luciferase coding sequence is ensured only after a precise ligation, and hence it reflects the precise NHEJ activity, while expression of active luciferase from a HindIII-linearized pGL vector in the luciferase promoter sequence requires any arbitrary ligation and thus reflects the overall NHEJ capacity. The luciferase activity was measured by a Dual Luciferase reporter assay (Promega) 48 h after transfection. The activity of pGL luciferase was normalized to a control circular pRL-tk vector.

2.8. Immunoprecipitation

The 293T cells were co-transfected with a pcDNA3.1-HA-wtBRCA or pcDNA3.1-HA-BRCA1Δ17–19 construct (8 μg) together with a pOZ-N-FLAG-Abraxas [26], pcDNA3.1-FLAG-CtIP or pcDNA3.1-FLAG-BARD1 construct (each 2 μg) using polyethylenimine and grown for 48 h. Subsequently the cells were extracted by an ice-cold EBC buffer (50 mM Tris pH 7.5, 150 mM NaCl, 0.5% NP-40, 1 mM EGTA) supplemented with 1 mM NaF and protease inhibitor cocktail (Roche), sonicated (3×10 s, amplitude 7, 100 W) and spun down (20,000 g, 10 min, 4 °C). Normalized cell extracts were incubated with an anti-FLAG M2 affinity gel (Sigma) for 3 h at 4 °C. Beads with immunoprecipitated proteins were extensively washed with an EBC buffer and then eluted with a sample buffer. Proteins were electrophoretically separated using 3–8% Tris-Acetate precast gradient gels (Invitrogen). The immunoblotted samples were then probed with anti-HA conjugated with HRP (Roche) or with anti-FLAG antibodies.

2.9. Statistical analysis

A statistical analysis was performed using the non-parametric ANOVA (Wilcoxon two sample test) test and an Excel spreadsheet. P values <0.05 were considered statistically significant. All data are presented as mean ± standard deviation (S.D.).

3. Results

3.1. BRCA1Δ17–19 sensitizes MCF-7 cells to mitomycin C

The MCF-7-based clones with a stably modified BRCA1 expression were selected and characterized following double-transfection with an expression construct containing the analyzed BRCA1 ASV and shRNA targeting wtBRCA1, using a procedure described previously [15]. For details see Supplementary Fig. 2.

BRCA1 is a key regulator of homology-directed DNA repair pathway(s) responsible for the DNA interstrand cross-link repair. Hence, depletion of BRCA1 sensitizes cells to DNA cross-linking agents such as mitomycin C, and the observed sensitivity of cells to mitomycin C reflects HR capacity, a major DNA repair mechanism operating in S/G2 cell-cycle phases of proliferating cells [1]. Considering this, we first examined the proliferation rate of the MCF-7 clones with a modified expression of BRCA1 treated by different concentrations of mitomycin C. Cell growth was continually monitored by an RTCA. The sensitivities of particular clones to mitomycin C were assessed by an estimate of the EC-50 concentration (Table 1) and evaluation of growth curves (Fig. 1). The MCF-7 clone with downregulated expression of endogenous BRCA1 (sh5196) was the one most sensitive to mitomycin C, responding rapidly to the addition of an active substance by slowing its growth rate. Both types of clones expressing the BRCA1Δ17–19 ASV [either (Δ17–19 + sh5196) alone or together with wtBRCA1 (Δ17–19)] also showed significantly increased sensitivity to mitomycin C compared to controls (MCF-7 and pcDNA3.1^{empty}). These results indicate that ectopically expressed BRCA1Δ17–19 overrides the capacity of homology-directed DNA repair in MCF-7 cells.

Table 1
The effective concentration (EC-50) of mitomycin C for MCF-7 clones with modified expression of BRCA1. The EC-50 was calculated at a time point of 64 h from the seeding of cells onto the plate as a cell index at a time point *versus* concentration. Data represent means from triplicates ± S.D.

Mitomycin C EC-50	MCF-7	Clone			
		pcDNA3.1	shRNA5196	Δ17–19	Δ17–19 + sh5196
[μg/ml] ± S.D.	6.18 ±0.18	6.06 ±0.23	2.89 ±0.27	4.02 ±0.71	3.60 ±0.2

3.2. BRCA1Δ17–19 slows down repair of ionizing radiation-induced DNA damage

Next, we directly measured the extent of DNA damage and DNA repair kinetics by a neutral comet assay in MCF-7 clones expressing the BRCA1Δ17–19 ASV following γ-irradiation (Fig. 2).

There were no detectable differences in DNA repair between the clone with downregulated wtBRCA1 expression (sh5196) and clones expressing the BRCA1Δ17–19 ASV. Notably however, the DNA repair velocity in cells with downregulated endogenous wtBRCA1 expression and cells expressing BRCA1Δ17–19 was slower compared to controls. The greatest difference in the degree of DNA damage between clones with modified expression of BRCA1 and controls was registered at 15 min post irradiation (PI). From that time on, the DNA damage rate progressively decreased, but at 120 min PI it was still significantly higher in clones with a modified BRCA1 expression.

These results show that both downregulation of wtBRCA1 and ectopic expression of the BRCA1Δ17–19 ASV slow down DDSB repairs in MCF-7 cells. Surprisingly, the expression of Δ17–19 negatively affected the DDSB repair capacity of MCF-7 cells similar to downregulation of endogenous wtBRCA1. The observed delay of the DDSB repair early after irradiation indicates that the BRCT protein interaction motif that is missing in the BRCA1Δ17–19 ASV is required primarily in the initial phase of DDSB repair, though it is not indispensable for the completion of the DNA repair process.

3.3. BRCA1Δ17–19 negatively impacts IRIF dynamics

The BRCA1 protein accumulates, along with many other DDR signaling and repair proteins, at DDSB-flanking regions of dynamically modified chromatin that are microscopically visible, commonly referred to as ionizing radiation-induced foci (IRIF), and eventually mediate repair of the DDSB lesions [27]. BRCA1 serves as a protein interaction modulator mediating the association of specific protein factors important for all phases of DNA repair. Thus, we further examined if the presence of the BRCA1Δ17–19 ASV that lacks a substantial part of the BRCA1 protein-interaction motif (in the first BRCT domain) can influence the kinetics of IRIF assembly and disassembly in response to ionizing radiation-induced DNA damage. For this purpose, we assessed the kinetics of γH2AX and 53BP1 co-localizations (well-established surrogate markers for IRIF) in fixed cells using immunofluorescence confocal microscopy after a single dose of 1.5 Gy of γ-radiation. Given that all analyzed clones originated from the MCF-7 cell line, and were exposed to the same radiation dose, it could be expected that identical initial average numbers of DDSBs occur in these settings.

The maximum number of IRIF in irradiated control MCF-7 cells was already detected at 5 min PI (Fig. 3). IRIF dissociation was apparent from that time on: over 75% of IRIF were dismantled within the first 2 h PI and the remaining IRIF subsequently disappeared reaching basal levels by 24 h PI, consistent with the expected kinetics of DNA repair. In MCF-7 clones with shRNA-mediated knock-down of endogenous wtBRCA1 (sh5196), and in clones expressing the BRCA1Δ17–19 ASV (Δ17–19 and Δ17–19 + sh5196, respectively), the number of IRIF further increased beyond the time of 5 min PI, reaching peak values at 30 min PI. The kinetics of the subsequent decline of the formed IRIF

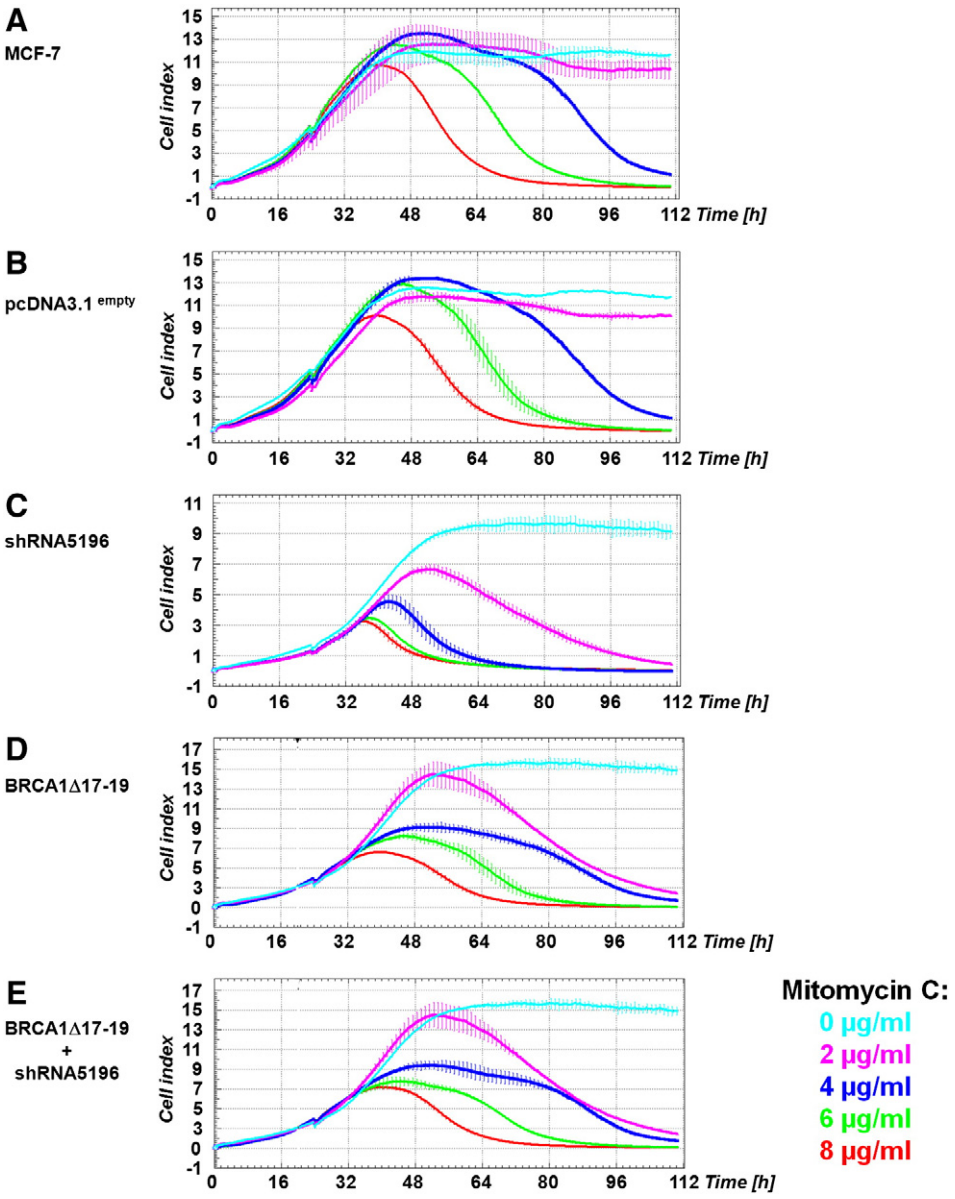


Fig. 1. Sensitivity of MCF-7 clones with modified expression of BRCA1 to mitomycin C. Cells were cultured for 5 days in a medium with different concentrations of mitomycin C (0, 2, 4, 6, and 8 µg/ml), under continual measurement of proliferation activity expressed here as a cell index (CI) using the xCELLigence RTCA analyzer. The CI was normalized to the time of addition of mitomycin C. Data are mean ± S.D.

were comparable to controls, but only 40% of IRIF (compared to 75% in control cells) were repaired by 2 h PI, likely reflecting the initial delay in IRIF formation and persistence. Interestingly, after 24 h PI the number of IRIF in MCF-7 clones with downregulated wtBRCA1 (sh5196) decreased to the basal pre-irradiation level, while in clones expressing BRCA1Δ17–19 IRIF numbers were still more than twice above the initial basal level (Fig. 3).

These findings are consistent with the results of comet assays and show that both downregulations of wtBRCA1 and ectopic expressions of the BRCA1Δ17–19 variant impair DDSB repair and the kinetics of IRIF formation. Surprisingly, downregulation of endogenous wtBRCA1 negatively influenced the assembly of IRIF while the decomposing phase of IRIF dynamics was unaffected. On the other hand, ectopic expression of the BRCA1Δ17–19 ASV impaired both phases of IRIF dynamics. Moreover, the negative impact on IRIF formation in the latter cells was apparently independent of the presence of wtBRCA1.

3.4. BRCA1Δ17–19 reduces NHEJ activity

BRCA1 is involved in both major DDSB repair pathways (HR and NHEJ) and negative effects of BRCA1 inactivation on NHEJ capacity [28] and fidelity [29] have been described. To find out whether ectopic expression of the BRCA1Δ17–19 ASV influences NHEJ activity and precision in MCF-7 cells, we performed an *in vitro* NHEJ repair assay. As can be seen from the data presented in Fig. 4, both the overall activity and the precision of NHEJ were significantly lower in all examined clones with altered BRCA1 expression, including clones with downregulated expression of wtBRCA1, and those expressing the ectopic BRCA1Δ17–19 variant. Notably, while the decrease in the overall NHEJ activity was similar in cells with knocked-down wtBRCA1 (to 34% ± 6.2 of the control values) and those expressing BRCA1Δ17–19 (a drop to 34% ± 3.8), the precision aspect of NHEJ was more dramatically affected in the latter clones: being decreased to 22% ± 3.2, as compared to a decline to 52% ± 3.4 seen upon shRNA-mediated knock-down of endogenous BRCA1 (Fig. 4).

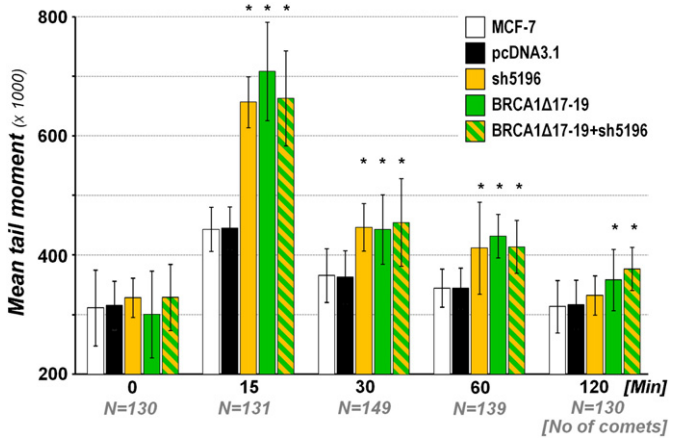


Fig. 2. DNA damage time-course in clones with modified BRCA1 expression after γ-irradiation. A neutral comet assay was used to determine the DNA damage level. At least 30 comets were analyzed for each clone using fluorescence microscopy at times 0, 15, 30, 60, and 120 min PI. For a DDSB repair time course quantification, the comet mean tail moment was calculated using the CometScore ver. 1.5 software (AutoComet). Non-irradiated samples are represented as time 0. Data are mean ± S.D. $^{*}p<0.01$ for MCF-7, (Wilcoxon test).

These results show that downregulation of wtBRCA1 as well as ectopic expression of the BRCA1Δ17–19 ASV negatively influence both the overall activity and fidelity of NHEJ in MCF-7 cells.

3.5. BRCA1Δ17–19 fails to interact with a subset of BRCA1 binding partners

Skipping exons 17–19 of BRCA1 leads to the production of a protein isoform missing a substantial part of the BRCT phosphoprotein-interaction module (aa 1664–1732). It can be expected that the protein-binding capacity of this BRCA1 isoform will be impaired. The inability to interact with specific binding partners could explain the results of previous experiments which showed an altered IRIF kinetics in MCF-7 clones expressing the BRCA1Δ17–19 ASV. To test the protein–protein interaction ability of the variant BRCA1 protein, we next analyzed the interactions between BRCA1Δ17–19 and two proteins, Abraxas and CtIP, that constitute a core of BRCA1-containing complex BRCA1-A and BRCA1-C, respectively. Both Abraxas and CtIP interact with BRCA1 *via* the BRCT domain and actively participate in DNA repair. Our results showed that in non-irradiated 293T cells, the BRCA1Δ17–19 isoform fails to co-precipitate with either Abraxas or CtIP, respectively, contrary to proficient interactions with either partner in case of wtBRCA1 (Fig. 5). To prove that the inability of BRCA1Δ17–19 to interact with Abraxas and CtIP is attributable to the truncation in the C-terminal BRCT domain, we further immunoprecipitated BRCA1Δ17–19 with BARD1 that binds to BRCA1 exclusively *via* the N-terminal RING domain. In this case, both the wtBRCA1 and the BRCA1Δ17–19 variant proteins co-precipitated with BARD1.

These results show that the BRCA1Δ17–19 ASV is unable to form stable protein complexes with important BRCA1 binding partners such as Abraxas and CtIP. This finding complements the above functional assays of IRIF kinetics and repair, and implies that the BRCA1Δ17–19 ASV protein is unable to execute the role of a protein-interaction modulator, normally provided by the wild-type BRCA1 protein, due to an impaired binding capacity of the C-terminal BRCT domain.

4. Discussion

In this study, we characterized the impact of the BRCA1Δ17–19 ASV on cell proliferation, sensitivity to DNA-damaging insults and capacity for DDSB repair. We identified this in-frame ASV previously in two *BRCA1* mutation-negative BC individuals during genetic screens of high-risk families [18]. Interestingly, intragenic *BRCA1* deletions affecting region coding for exons 17–19 that result in expression of identical BRCA1 protein isoform were reported in high-risk BC patients of the Western European descent [30,31]. As discussed below, this work extends our previous report indicating the importance of BRCA1 ASVs for the maintenance of genome integrity [15], and provides novel insights into the biological impact and mechanistic basis for malfunction of the so far uncharacterized variant BRCA1Δ17–19 of the BRCA1 tumor suppressor protein.

Timely and error-free repair of a DDSB in eukaryotic cells is secured by a cooperative action of NHEJ and HR pathways [32]. NHEJ is the prominent repair mechanism during the G1 and early S phase, and throughout the cell cycle when DDSBs occur in easily accessible euchromatin [33]. Due to low demands for the modification of free DNA ends at the breakage site, NHEJ is a relatively simple mechanism and represents a fast component of the global DDSB repair that ensures prompt elimination of such highly deleterious DNA lesion [34]. The exact participation of BRCA1 in NHEJ is not understood yet, and the results concerning BRCA1's importance for NHEJ are contradictory [35,36]. The *in vivo* regulation of NHEJ relies on extensive chromatin modifications (methylation, ubiquitination) that are unlikely to be faithfully recapitulated in ectopically expressed constructs used as reporters for cell culture analyses [37]. Hence, the results of such reporter-based NHEJ assay may not exactly reflect the activity of NHEJ *in vivo*. Such concerns notwithstanding, our results indicate that downregulation of endogenous wtBRCA1 or ectopic expression of the BRCA1Δ17–19 variant has a negative impact on the capacity of NHEJ *in vitro*. This is consistent with the observed delay of DNA repair in the period shortly after DDSB induction.

Homology-directed DNA repair represents a slow but precise component of DDSB repair ensuring fidelity in elimination of a highly deleterious DNA lesion. Our present confocal immunofluorescence microscopy analysis showed that the formation of γH2AX/53BP1 IRIF was significantly retarded within 30 min PI and further persistence of IRIF was markedly prolonged in MCF-7 cells pretreated by either shRNA-mediated knockdown of wtBRCA1 or ectopic expression of BRCA1Δ17–19 ASV. These results, together with increased sensitivity to mitomycin C, document a negative impact of the analyzed variant BRCA1 protein expression on HR.

The BRCA1 protein participates in several protein complexes that ensure discrete successive steps of HR [4]. The inability of BRCA1 to interact with its binding partners results in delayed and/or erroneous spatio-temporal regulation and function at the sites of DNA damage which, in turn, deregulate the highly orchestrated DDSB repair process. Indeed, a delay of a single rate-limiting step can delay the entire repair process, and a total blockage in one of the HR steps may cause a

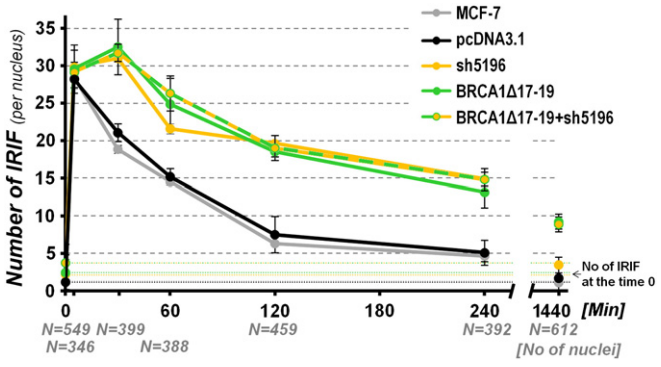


Fig. 3. The kinetics of IRIF formation and persistence in diverse clones of MCF-7 cells. The kinetics of DDSB signaling/repair were assessed by counting the γH2AX/53BP1 colocalizations per nucleus at the indicated time points (0, 5, 30, 60, 120, 240 and 1440 min, respectively) after ionizing irradiation (1.5 Gy) using confocal immunofluorescence microscopy. The endogenous level of DDSBs in the indicated clones is represented here as the number of IRIF per nucleus in non-irradiated cells (time 0; depicted by dotted lines). Data are mean ± S.D.

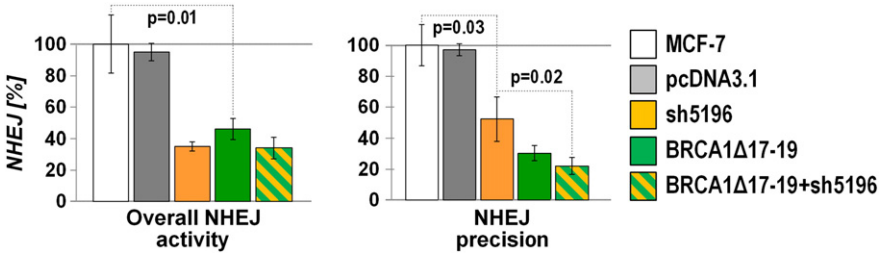


Fig. 4. The *in vitro* NHEJ assay. The particular clones and relevant controls were co-transfected by a linearized pGL-control vector and a circular pRL-tk vector. The rescued activity of luciferase was measured using the dual luciferase assay kit and normalized to the activity of control pRL-tk luciferase. The luciferase activity rescued from the HindIII-linearized pGL-control vector reflects the overall NHEJ activity, while that from the EcoRI-linearized pGL-control vector reflects the NHEJ precision [15]. The experiment was repeated three times, data are mean ± S.D., significant p-values are indicated.

time-consuming functional switch redirecting the DDSB repair from HR to another repair pathway [38].

The formation of long 3′ ssDNA overhangs of the free DNA ends at the DDSB site enables a successful strand exchange for a homology-directed DNA repair [9]. These overhangs are dominantly generated by the activity of CtIP exonuclease. The rate of the CtIP-mediated DNA end resection increases as CtIP is bound to BRCA1 within the BRCA1 C-complex, while the activity of unbound CtIP molecules is inhibited by their association with the DNA repair modulator 53BP1 at the DDSB site [39,40]. Limited resection of DNA ends is insufficient for further HR, and therefore the integration of CtIP into the BRCA1 C-complex overcomes the inhibitory activity of 53BP1 and enables the formation of long 3′ DNA overhangs [41]. Impaired resection caused by the inability of the BRCA1Δ17–19 isoform to bind CtIP can cause the observed delay in DDSB repair. Defective assembly of the BRCA1 C-complex in MCF-7 cell clones expressing the BRCA1Δ17–19 isoform as well as in those with downregulated wtBRCA1 expression can explain their similar behavior in terms of increased sensitivity to mitomycin C and delayed early phase of DNA repair.

Once the DDSBs are fully repaired, HR must be resolved by dismantling the active IRIFs to prevent a potential undesirable DNA crossover (referred to as hyper-HR) [2]. Taking into account the results of comet assays showing that the majority of DDSBs are repaired within 120 min PI and colocalization studies showing increased persistence of IRIFs beyond this time, it could be suggested that the presence of the BRCA1Δ17–19 variant interferes with processing of the slowly-repaired DDSBs localized in the highly complex chromatin regions [42] or disturbs the termination of HR by blocking the IRIF removal.

The consecutive association of protein factors that form the active IRIF at the DDSB site is regulated by covalent protein modifications. Besides phosphorylation, the importance of ubiquitination is indicated by recent findings showing that ubiquitination and deubiquitination of histones are critical for genome stability maintenance [27,43]. The BRCA1 ubiquitin ligase activity, enhanced by heterodimerization with BARD1 *via* the RING domain-mediated interaction, targets several proteins directly involved in the DNA repair process [44,45]. Alongside the E3 activity of BRCA1, the BRCA1 A-complex anchoring to the DDSB site binds BRCC36 deubiquitinylase [46]. Though the exact mechanism of the IRIF disassembly is currently unknown, Hu et al. reported that downregulation of non-BRCA1 members of the BRCA1 A-complex causes hyper-HR [47]. Furthermore, Dever et al. showed that BRCA1 mutation in its BRCT domain causes increased recombination leading to genomic instability [48]. These results indicate that the BRCA1 A-complex is a negative regulator of ubiquitination-dependent DNA repair pathways. The inability of the BRCA1Δ17–19 ASV to bind Abraxas, a central protein of the BRCA1 A-complex, might contribute to the observed prolonged persistence of IRIF. Moreover, the BRCA1 mutation p.I26A affecting the RING domain was shown to reverse the increased recombination rate in cells expressing the BRCA1 BRCT mutant p.K1702M [48]. This is consistent with our observation that persistence of IRIF was prolonged in cell ectopic expression BRCA1Δ17–19 (which maintains the RING domain-mediated BARD1 binding) but not in cells with downregulated wtBRCA1.

More than twenty BRCA1 ASVs have been described so far [13,14]. The majority of them are classified as variants of unknown clinical significance. However, several of these ASVs were shown to negatively

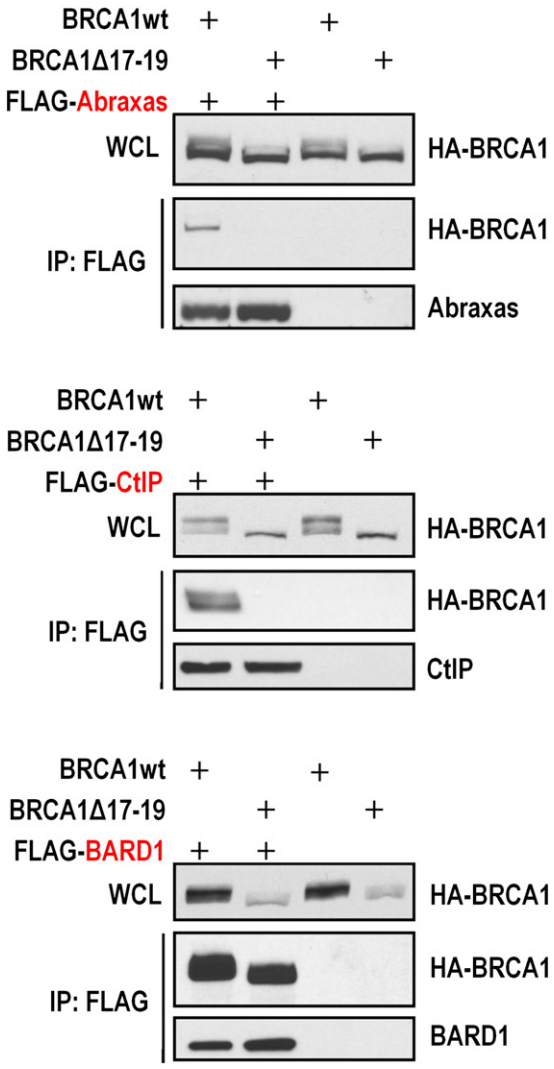


Fig. 5. Immunoprecipitation (IP) of the BRCA1Δ17–19 ASV with Abraxas, CtIP, and BARD1, respectively. 293T cells were transiently co-transfected with an expression vector containing HA-tagged full-length BRCA1 or HA-tagged BRCA1Δ17–19 variant together with vectors containing FLAG-tagged BARD1, Abraxas or CtIP, respectively. The expression levels of the individual proteins were validated by western blotting (WB) in whole-cell lysates (WCL). Co-immunoprecipitation was performed in non-irradiated cells. The results show that the Δ17–19 isoform with the deletion of a substantial part of the BRCT domain is unable to interact with the natural BRCA1 C-terminal binding partners Abraxas and CtIP, while the intact N-terminal RING domain-mediated binding to BARD1 is preserved.

influence the BRCA1-mediated cell cycle control and DNA repair activity [49,15]. This indicates that BRCA1 ASVs generated in a tissue-specific manner by misregulated pre-mRNA splicing can also contribute to tumorigenesis. It has been observed that the formation of BRCA1 aberrant splicing variants can promote malignant transformation [50–52]. On the other hand, the involvement of regulated AS in cancer development is virtually unknown [53]. For numerous gene products, including BRCA1, it has been shown that the formation of specific ASVs can be determined by a cell cycle phase [54], or by the DNA damage response [55,56]. This suggests that AS is an important post-transcriptional regulatory event responding to specific signals. Thus AS misregulation can lead to formation of an alternative protein product with potentially aberrant function or uncoupled from upstream regulation. The resulting appearance of ASVs that should normally be created under different cellular conditions and/or in a different amount can markedly alter the downstream processes that depend on the protein subjected to AS, with potential implications for human diseases including cancer.

In our previous study we showed that ectopic expression of the BRCA1Δ14–15 ASV that lacks a part of the phosphorylation-regulated SCD does not change the sensitivity of MCF-7 cells to mitomycin C but impairs the DDSB repair capacity by decreasing the NHEJ activity [15]. Our current results show a distinct, genome-destabilizing phenotype, in that ectopic expression of the BRCA1Δ17–19 ASV in MCF-7 cells significantly increases their sensitivity to mitomycin C. This indicates a negative effect of the BRCA1Δ17–19 ASV on DDSB repairs with a broader impact on the HR pathway, with potential implications for tumorigenesis and sensitivity to emerging new treatments such as PARP inhibitors, that selectively target cancer cells with functionally impaired homologous recombination [57].

5. Conclusions

Based on the present results we conclude that the ectopically expressed BRCA1Δ17–19 ASV protein with an in-frame deletion affecting a substantial part of the BRCT domain is unable to interact with Abraxas and CtIP while it retains the BARD1-binding capacity. This defect probably causes the observed hypersensitivity to mitomycin C and the reduced activity of ionizing radiation-induced DDSB repair in MCF-7 cells. The presence of the BRCA1Δ17–19 variant decelerates DNA repair and the assembly of IRIF during early PI periods, and impairs the NHEJ activity. Similar phenotypes were observed in cells with downregulated expression of endogenous wtBRCA1, indicating that the Δ17–19 ASV is unable to functionally substitute full-length BRCA1 in DNA repair. Moreover, expression of the BRCA1Δ17–19 isoform significantly prolongs the persistence of IRIF regardless of the actual DNA damage level. In a broader context, our results suggest that expression of in-frame ASVs of BRCA1 mRNA that selectively lack portions of critical structural domains may negatively affect genome integrity by undermining the overall capacity or precision of DNA repair, and possibly promoting hyperactivity of HR.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.cellsig.2013.02.008.

Acknowledgment

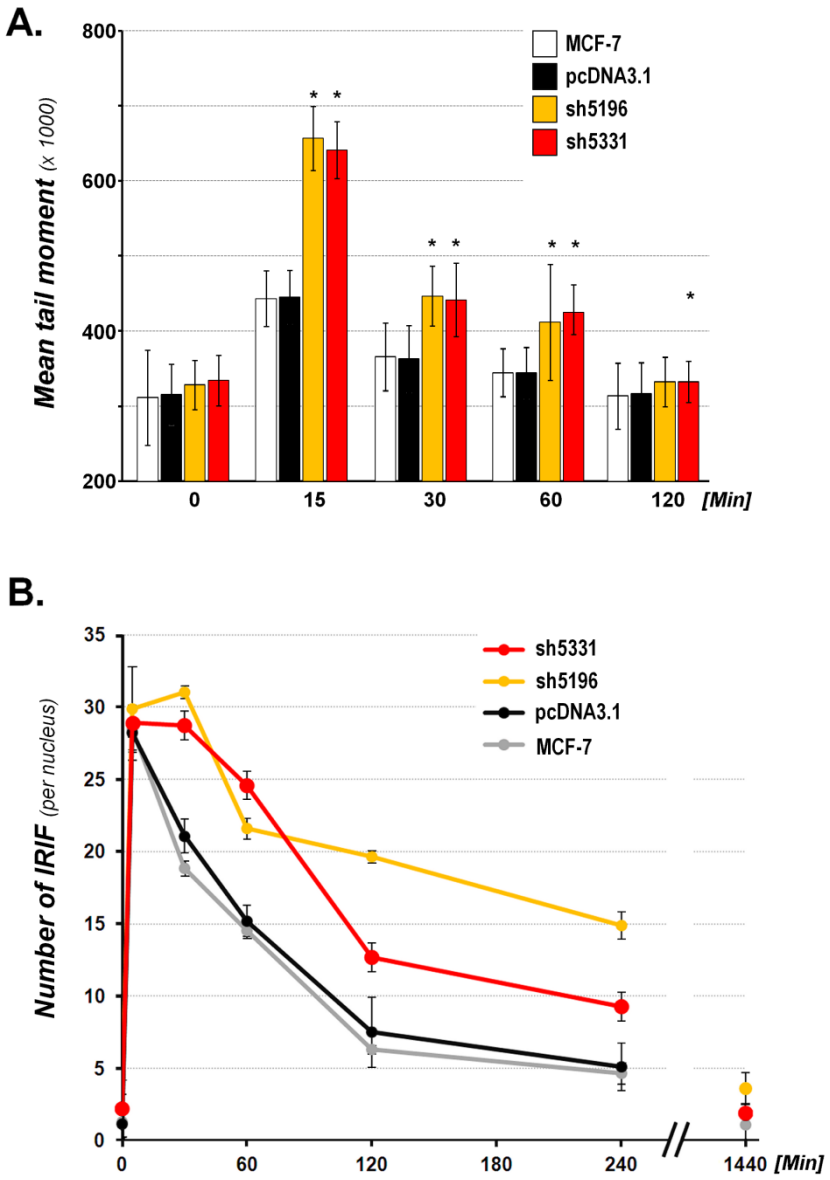
We thank all members of our laboratories and Jan Flemr for proofreading. This study was supported by grants from the Grant Agency of the Czech Republic (No. P301/12/1850 and P301/10/1525), the Internal Grant Agency of the Ministry of Health (No. NT12280 and NT11065), the European Commission (projects DDResponse and Biomedreg), the Charles University Grant Agency (No. 428711), MEYS COST LD12039, GACR Center Excellence (302/12/G157), and the Charles University in Prague (project PRVOUK-P27/LF1/1).

References

[1] M.E. Moynahan, T.Y. Cui, M. Jasin, Cancer Research 61 (2001) 4842–4850.
[2] M. Jasin, Oncogene 21 (2002) 8981–8993.
[3] S.M. Sy, M.S. Huen, J. Chen, Proceedings of the National Academy of Sciences of the United States of America 106 (2009) 7155–7160.
[4] B. Wang, Cell & Bioscience 2 (2012) 6–12.
[5] R.A. Greenberg, B. Sobhian, S. Pathania, S.B. Cantor, Y. Nakatani, D.M. Livingston, Genes & Development 20 (2006) 34–46.
[6] L. Feng, J. Wang, J. Chen, Journal of Biological Chemistry 285 (2010) 30982–30988.
[7] S. Panier, D. Durocher, DNA Repair (Amst) 8 (2009) 436–443.
[8] L. Dohrn, D. Salles, S.Y. Siehler, J. Kaufmann, L. Wiesmuller, Biochemical Journal 441 (2012) 919–926.
[9] A.A. Sartori, C. Lukas, J. Coates, M. Mistrik, S. Fu, J. Bartek, R. Baer, J. Lukas, S.P. Jackson, Nature 450 (2007) 509–514.
[10] M.T. Mok, B.R. Henderson, Traffic 13 (2012) 800–814.
[11] B.R. Henderson, Bioessays 27 (2005) 884–893.
[12] W.D. Foulkes, S.A. Narod, Clinical and Investigative Medicine 18 (1995) 473–483.
[13] T.I. Orban, E. Olah, Molecular Pathology 56 (2003) 191–197.
[14] C.A. Pettigrew, J.D. French, J.M. Saunus, S.L. Edwards, A.V. Sauer, C.E. Smart, T. Lundstrom, C. Wiesner, A.B. Spurdle, J.A. Rothnagel, M.A. Brown, Breast Cancer Research and Treatment 119 (2010) 239–247.
[15] J. Sevcik, M. Falk, P. Kleiblova, F. Lhota, L. Stefancikova, M. Janatova, L. Weiterova, E. Lukasova, S. Kozubek, P. Pohlreich, Z. Kleibl, Cellular Signalling 24 (2012) 1023–1030.
[16] M. Takahashi, Y. Furukawa, H. Shimodaira, M. Sakayori, T. Moriya, Y. Moriya, Y. Nakamura, C. Ishioka, Familial Cancer 11 (2012) 559–564.
[17] L. Schwarzova, J. Stekrova, M. Florianova, A. Novotny, M. Schneiderova, P. Lnenicka, V. Kebrdlova, J. Kotlas, K. Vesela, M. Kohoutova, Familial Cancer (2012), http://dx.doi.org/10.1007/s10689-012-9569-8.
[18] P. Pohlreich, J. Stribrna, Z. Kleibl, M. Zikan, R. Kalbacova, L. Petruzelka, B. Konopasek, Medical Principles and Practice 12 (2003) 23–29.
[19] J.E. Quinn, R.D. Kennedy, P.B. Mullan, P.M. Gilmore, M. Carty, P.G. Johnston, D.P. Harkin, Cancer Research 63 (2003) 6221–6228.
[20] E. Vondruskova, R. Malik, J. Sevcik, P. Kleiblova, Z. Kleibl, Neoplasma 55 (2008) 130–137.
[21] C. Ehrhardt, M. Schmolke, A. Matzke, A. Knoblauch, C. Will, V. Wixler, S. Ludwig, Signal Transduction 6 (2006) 179–184.
[22] M. Falk, E. Lukasova, S. Kozubek, Biochimica et Biophysica Acta 1783 (2008) 2398–2414.
[23] M. Falk, E. Lukasova, B. Gabrielova, V. Ondrej, S. Kozubek, Biochimica et Biophysica Acta 1773 (2007) 1534–1545.
[24] M. Kozubek, S. Kozubek, E. Lukasova, E. Bartova, M. Skalnikova, P. Matula, P. Matula, P. Jirsova, A. Cafourkova, I. Koutna, Cytometry 45 (2001) 1–12.
[25] A. Sharma, K. Singh, A. Almasan, Methods in Molecular Biology 920 (2012) 613–626, http://dx.doi.org/10.1007/978-1-61779-998-3_40, (613-626).
[26] J. Patterson-Fortin, G. Shao, H. Bretscher, T.E. Messick, R.A. Greenberg, Journal of Biological Chemistry 285 (2010) 30971–30981.
[27] J. Lukas, C. Lukas, J. Bartek, Nature Cell Biology 13 (2011) 1161–1169.
[28] Q. Zhong, T.G. Boyer, P.L. Chen, W.H. Lee, Cancer Research 62 (2002) 3966–3970.
[29] E.G. Thompson, H. Fares, K. Dixon, Environmental and Molecular Mutagenesis 53 (2012) 32–43.
[30] M.A. Unger, K.L. Nathanson, K. Calzone, D. Antin-Ozerkis, H.A. Shih, A.M. Martin, G.M. Lenoir, S. Mazoyer, B.L. Weber, American Journal of Human Genetics 67 (2000) 841–850.
[31] T.V.O. Hansen, L. Jønson, A. Albrechtsen, M.K. Andersen, B. Ejlertsen, F.C. Nielsen, Breast Cancer Research and Treatment 115 (2009) 315–323.
[32] E.M. Kass, M. Jasin, FEBS Letters 584 (2010) 3703–3708.
[33] A.A. Goodarzi, P. Jeggo, M. Lobrich, DNA Repair (Amst) 9 (2010) 1273–1282.
[34] Y. Lorat, S. Schanz, N. Schuler, G. Wennemuth, C. Rube, C.E. Rube, PLoS One 7 (2012) e38165.
[35] M.E. Moynahan, J.W. Chiu, B.H. Koller, M. Jasin, Molecular Cell 4 (1999) 511–518.
[36] H.C. Wang, W.C. Chou, S.Y. Shieh, C.Y. Shen, Cancer Research 66 (2006) 1391–1400.
[37] A.T. Noon, A.A. Goodarzi, DNA Repair (Amst) 10 (2011) 1071–1076.
[38] A. Bothmer, D.F. Robbiani, M. Di Virgilio, S.F. Bunting, I.A. Klein, N. Feldhahn, J. Barlow, H.T. Chen, D. Bosque, E. Callen, A. Nussenzweig, M.C. Nussenzweig, Molecular Cell 42 (2011) 319–329.
[39] B. Sobhian, G. Shao, D.R. Lilli, A.C. Culhane, L.A. Moreau, B. Xia, D.M. Livingston, R.A. Greenberg, Science 316 (2007) 1198–1202.
[40] L. Chen, C.J. Nievera, A.Y. Lee, X. Wu, Journal of Biological Chemistry 283 (2008) 7713–7720.
[41] S.F. Bunting, E. Callen, N. Wong, H.T. Chen, F. Polato, A. Gunn, A. Bothmer, N. Feldhahn, O. Fernandez-Capetillo, L. Cao, X. Xu, C.X. Deng, T. Finkel, M. Nussenzweig, J.M. Stark, A. Nussenzweig, Cell 141 (2010) 243–254.
[42] L. Ježková, M. Falk, I. Falková, M. Davídková, A. Bačíková, L. Stefančíková, J. Vachelová, A. Michaelidesová, E. Lukášová, A. Boreyko, E. Krasavin, S. Kozubek, Applied Radiation and Isotopes (in press), http://dx.doi.org/10.1016/j.apradiso.2013.01.022.
[43] G. Shao, D.R. Lilli, J. Patterson-Fortin, K.A. Coleman, D.E. Morrissey, R.A. Greenberg, Proceedings of the National Academy of Sciences of the United States of America 106 (2009) 3166–3171.
[44] R. Scully, J. Chen, R.L. Ochs, K. Keegan, M. Hoekstra, J. Feunteun, D.M. Livingston, Cell 90 (1997) 425–435.
[45] G.T. Lok, S.M. Sy, S.S. Dong, Y.P. Ching, S.W. Tsao, T.M. Thomson, M.S. Huen, Nucleic Acids Research 40 (2012) 196–205.

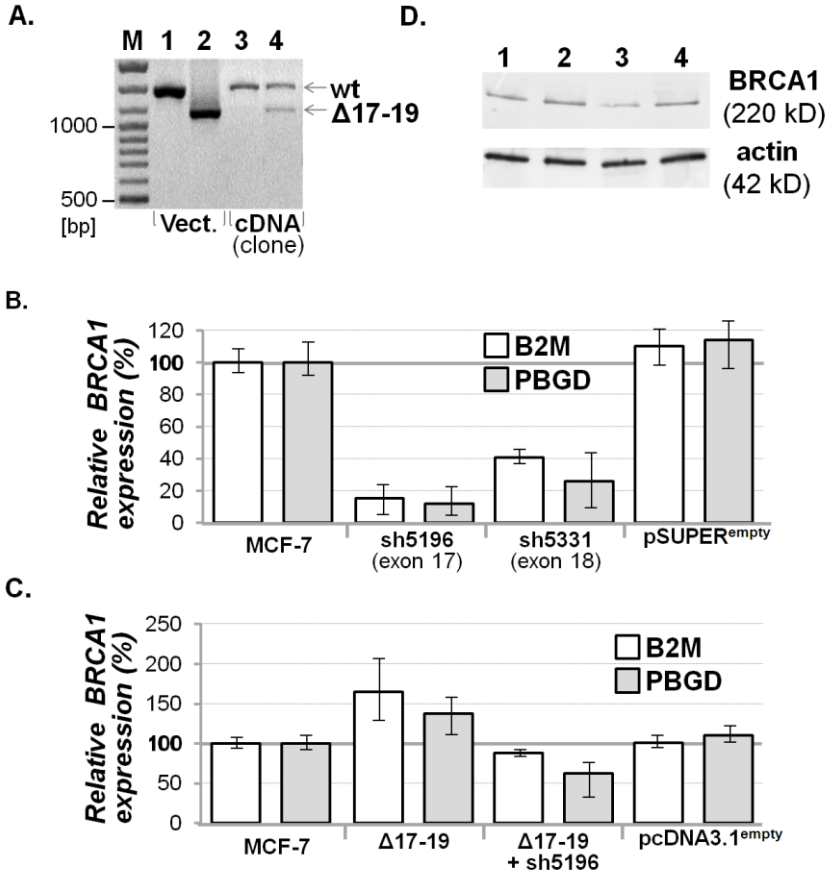
[46] Y. Dong, M.A. Hakimi, X. Chen, E. Kumaraswamy, N.S. Cooch, A.K. Godwin, R. Shiekhattar, Molecular Cell 12 (2003) 1087–1099.
[47] Y. Hu, R. Scully, B. Sobhian, A. Xie, E. Shestakova, D.M. Livingston, Genes & Development 25 (2011) 685–700.
[48] S.M. Dever, S.E. Golding, E. Rosenberg, B.R. Adams, M.O. Idowu, J.M. Quillin, N. Valerie, B. Xu, L.F. Povirk, K. Valerie, Aging (Albany, NY) 3 (2011) 515–532.
[49] K. Chock, J.M. Allison, W.M. Elshamy, Oncogene 29 (2010) 5274–5285.
[50] R.D. Brandao, K.E. van Roozendaal, D. Tserpelis, B. Caanen, G.E. Gomez, M.J. Blok, Breast Cancer Research and Treatment 131 (2012) 723–725.
[51] L. Zhang, L. Chen, R. Bacares, J.M. Ruggeri, J. Somar, Y. Kemel, Z.K. Stadler, K. Offit, Breast Cancer Research and Treatment 130 (2011) 1051–1056.
[52] M. Thomassen, A. Blanco, M. Montagna, T.V. Hansen, I.S. Pedersen, S. Gutierrez-Enriquez, M. Menendez, L. Fachal, M. Santamarina, A.Y. Steffensen, L. Jonson, S. Agata, P. Whiley, S. Tognazzo, E. Tornero, U.B. Jensen, J. Balmana, T.A. Kruse, D.E. Goldgar, C. Lazaro, O. Diez, A.B. Spurdle, A. Vega, Breast Cancer Research and Treatment 132 (2012) 1009–1023.
[53] E. Scholzova, R. Malik, J. Sevcik, Z. Kleibl, Cancer Letters 246 (2007) 12–23.
[54] T.I. Orban, E. Olah, Biochemical and Biophysical Research Communications 280 (2001) 32–38.
[55] M.J. Munoz, M.S. Perez Santangelo, M.P. Paronetto, M. de la Mata, F. Pelisch, S. Boireau, K. Glover-Cutter, C. Ben-Dov, M. Blaustein, J.J. Lozano, G. Bird, D. Bentley, E. Bertrand, A.R. Kornblihtt, Cell 137 (2009) 708–720.
[56] D.S. Chandler, R.K. Singh, L.C. Caldwell, J.L. Bitler, G. Lozano, Cancer Research 66 (2006) 9502–9508.
[57] S.P. Jackson, J. Bartek, Nature 461 (2009) 1071–1078.

Supplementary figure 1:



Supplementary figure 1. Comet assays (A) and high-resolution confocal microscopy (B) experiments were performed with two sets of clones expressing shRNAs [sh5196 (targeting BRCA1 exon 17) and sh5331 (targeting BRCA1 exon 18)] to rule out the off-target effect of the designed shRNAs. Data represent means from analyses of two clones (each in triplicate) ± S.D. Both shRNAs show similar effects. Less apparent impairment of sh5331 on IRIF disassembly is probably attributable to the weaker downregulation of wtBRCA1 reached by this shRNA (Supplementary Fig. 2B)

Supplementary figure 2:



Supplementary figure 2. The pcDNA3.1 BRCA1Δ17-19 expression construct was prepared by a PCR splicing approach. The shRNA sequences were designed according to generally accepted rules for a functional interfering RNA. The indicated shRNAs were targeted to the DNA regions missing in Δ17-19: sh5196 (exon 17) and sh5331 (exon 18), respectively. Stable MCF-7-derived clones were prepared by calcium-phosphate transfection followed by selection with hygromycin B and puromycin.

The functionality of the expression system was proved by RT-PCR (**A**; M – marker; 1 – pcDNA3.1 wtBRCA1 vector; 2 – pcDNA3.1 BRCA1Δ17-19 vector; 3 – cDNA prepared from non-transfected MCF-7 cells; 4 – cDNA prepared from MCF-7 cells transfected by a pcDNA3.1 BRCA1Δ17-19) and qPCR [**B, C** – relative expression of BRCA1 was normalized using β- 2 microglobulin (B2M) and (PBGD) as housekeeping] on the mRNA level and by western blotting (**D**; 1 . MCF-7 non-transfected; 2 – pcDNA3.1.^{empty}; 3 - sh5196, 4 - Δ17-19 + sh5196) on the protein level.