Bioinformatics and Molecular Analysis of the Breast Cancer Susceptibility Gene BRCA1 in Breast Cancer

Inam Jasim Lafta

Department of Microbiology, College of Veterinary Medicine, University of Baghdad, Baghdad, Iraq

Abstract

Background: The breast cancer susceptibility gene (BRCA1) encodes a tumor suppressor protein which plays a vital role in the DNA damage repair and transcriptional regulation among other functions. Bioinformatics is a newly-emerged discipline that uses computer, mathematics, and statistics in molecular biology in order to analyze the large amounts of biological data quickly, freely, and accurately.

Methods: The BRCA1 transcript (mRNA) levels were checked by using real-time quantitative polymerase chain reaction (RT-qPCR) in four sporadic breast cancer cell lines: MCF-7, T47D, MDA-MB-231, and MDA-MB-468 compared to the normal breast tissue. Bioinformatics tools were also used to compare and analyze different aspects of BRCA1 transcripts (multiple different mRNAs produced by a single gene) and splice variants (multiple proteins encoded by the same gene).

Results: The level of BRCA1 mRNA was overexpressed in the studied breast cancer cell lines relative to the normal breast cDNA. Also, the bioinformatics software tools provided many important features of this gene that would help explain numerous controversial laboratory findings.

Conclusions: The data presented here support a role for BRCA1 overexpression in the pathogenesis of sporadic breast cancer. The bioinformatics analysis of BRCA1 mRNA and protein variants can provide information essential for cancer diagnosis or therapy. The biological data gained from these tools can help authors make better decisions before launching expensive experiments.

Keywords: Breast cancer, BRCA1, Tumor suppressor gene, Gene expression, Bioinformatics

♦**Corresponding Author:**
Inam J. Lafta, PhD
Department of Microbiology,
College of Veterinary Medicine,
University of Baghdad,
Baghdad, Iraq
Tel: +964 7735869299
Email: ijlafta@gmail.com

Introduction

It has been confirmed that the breast cancer susceptibility gene (BRCA1) has an important role in the pathogenesis of sporadic and inherited breast cancer.[1] BRCA1 is a well-known breast and ovarian tumor suppressor gene encoding multiple proteins,[2] particularly p220, 11b, and IRIS.[3] Although the proteins 11b and p220 share sequences, most of the sequences encoded by the

largest exon of p220 (exon 11) are unavailable from the 11b protein sequence.[4] Normally, the BRCA1 protein is in a nuclear multi-protein complex including BRCA2, RAD51, and BARD1 and the DNA damage repair proteins including MLH1, MSH2, MSH6, ATM, MRE11, RAD50, NBS1, BLM, and RFC.[5]

The BRCA1 protein is essential in the maintenance of genomic stability through many mechanisms. It is involved in transcriptional regulation through interaction with several transcription factors,[6] control of DNA repair,[5] and participation in various signaling pathways implicated in transcription and checkpoint control.[7] Also, it participates in growth suppression, ubiquitin ligation and apoptosis induction.[8] BRCA1 is likely to be implicated in biosynthesis of ribosomes, progression of the cell cycle, and maintenance of complexes created in response to DNA repair and cellular stress.[9,10] The best known of these functions, that is double repaired by homologous recombination (HR), is the double-stranded DNA (dsDNA),[4] through which BRCA1 is supposed to be implicated in the breast cancer suppression.[11-13] Moreover, BRCA1 involves in the repair of stalled or collapsed replication forks,[14,15] assists in FANCD2 localization during the repair of inter-strand cross-link,[16-18] formation of mitotic spindle pole,[19] maintenance of normal centrosome number,[20,21] the inhibition of satellite RNA expression,[22] suppression of translesional synthesis, and base mutagenesis.[14,15]

The function of BRCA1 in tumorigenesis has been found to be complicated and the decreased expression of BRCA1 can result in numerous abnormalities, including HR pathway defect due to hypersensitivity to several agents that cause dsDNA breaks (e.g. ionizing radiation).[1] Almost all of the above functions are attributed to p220 (also known as BRCA1), which is a 220 kDa multifunctional nuclear phosphoprotein.[23] Little is known regarding the function of 11b and IRIS as well.[4] However, it is believed that IRIS is an endogenous protein that normally motivates DNA replication and can regulate specific transcriptional events by adding to the exhibition of oncoprotein properties upon endogenous overexpression.[24,25] The BRCA1-IRIS (also called "IRIS"), is an alternatively spliced BRCA1 protein and a chromatin-bound replication and transcription regulator, which is overexpressed (i.e. oncoprotein) in many carcinomas, including breast cancer.[2] Unfortunately, few studies, if not at all, focused on the overexpressed BRCA1 (with malfunctions) in breast cancer, within which it might drive a role similar to that of mutated BRCA1.

Comparisons between multiple BRCA1 transcripts (numerous different mRNAs produced by a single gene) and isoforms (multiple proteins encoded by a single gene) in terms of molecular structure, chemical and biological features by means of freely available bioinformatics tools could save time, labor, and money leading to further insights into the theoretical basis for the molecular pathways involved in BRCA1 protein function and probably the further development of targeted therapies.[1] Bioinformatics is defined as the use of computer tools to analyze and interpret biological data.[26] It is an interdisciplinary area that exploits computer science, mathematics, biology, and physics and is necessary for managing data in medicine and contemporary biology.[27] Therefore, the present study analyzed the mRNA expression of BRCA1 in breast cancer cell lines and the normal breast tissue by using the real time quantitative polymerase chain reaction (RT-qPCR) along with the gene and its products by the bioinformatics. This analysis encompasses the comparison of the sequences of BRCA1 transcript variants and isoforms, physical and chemical properties, secondary and three-dimensional (3-D) structures, and phylogenetic relationships in order to provide furthermore information and some explanations to numerous controversial laboratory data.

Materials and Methods
Cell culture
The breast cancer cell lines used in this study included MCF-7, T47D, MDA-MB-231, and MDA-MB-468 which were purchased from Sigma-Aldrich, UK. All of these cells were grown

in Dulbecco's modified eagles medium (DMEM; Lonza) containing 4.5 g/L glucose with L-glutamine, and supplemented with 10% fetal calf serum (FCS; Seralab) and 1× non-essential amino acids (NEAAs; Bio Whittaker). Before use, the medium, Trypsin-Versene EDTA (Lonza), and phosphate-buffered saline (PBS) were warmed in a 37˚C water bath for at least 30 minutes.

RNA Extraction and conversion into cDNA

To analyze the BRCA1 mRNA expression in the breast cancer cell lines whose cell pellets were collected, total cell RNA was extracted by using mammalian total RNA miniprep kit (Sigma-Aldrich). Manufacturer's instructions were followed and finally the RNA was eluted in 50µl RNase free water. Subsequently, the RNA concentration was measured using NanoDrop spectrophotometer (Thermo Fisher Scientific), and it was converted immediately to cDNA. The reverse transcription reaction to convert RNA to cDNA was performed following the instructions of the manufacturer of a High Capacity RNA to cDNA Kit (AB Applied Biosystems). The total volume of the cDNA (20µl) was kept at -20˚C until use. Concerning the whole RNA of the normal breast tissue obtained commercially, it was converted to cDNA by using the same kit mentioned above.

RT-qPCR

The expression of BRCA1 and 18S was studied by RT-qPCR using a Corbett Robotics Rotor-Gene™ 6000 (Qiagen). Primers specific for BRCA1 and 18S (Eurofins, Germany) were used in RT-qPCR which included BRCA1 forward 5'-CCCTTCACCAACATGCCCACA-3' and BRCA1 reverse 5'-CTGCCCAATTGCATG-GAAGCC-3', and for 18S the primers were forward 5'-AGAAACGGCTACCACATCCA-3' and reverse 5'-CACCAGACTTGCCCTCCA-3'. They were checked for their specificity by using the National Center for Biotechnology Information (NCBI) bioinformatics website, and BRCA1 primers were designed to be able to amplify all of the variants. The PCR reaction consisted of 20 µl of: 2× SensiMix (10 µl) (a mixture of buffer,

dNTP, SYBR Green, HiRox and modified Taq polymerase). In addition, the reaction contained primers (10 µM each in 2 µl of forward and reverse), cDNA (1µg; 5 µl) as a template and deionized sterile water (3 µl). Reactions were carried out in triplicate technical repetitions. Most importantly, melt curve analysis was applied to check for the purity of the amplified products and the absence of primer dimmers. The cycling conditions involved: 10 min at 95°C for Taq polymerase activation followed by 40 cycles of denaturation at 95°C for 15 s, annealing at 58°C for 15 s and extension at 72°C for 30 s. The CT values were determined for each cDNA sample in every reaction by means of Rotor-Gene 6000 software. The amplification efficiency was calculated for the reference and target genes, and the average was equal to 1.9. The fold change in gene expression was calculated according to the formula $1.9^{-\Delta\Delta CT}$ to compare the expression level relative to the expression of same gene in the normal breast tissue. Furthermore, $\Delta\Delta$ CT was measured based on the formula: Δ CT (cancer) - Δ CT (normal), while Δ CT = CT (target gene) - CT (reference gene).[28]

Bioinformatics analysis of BRCA1 transcripts and isoforms

Using the online NCBI and Ensembl software tools, the homosapiens BRCA1 gene, its size, transcripts, and isoforms were searched. Diagrams depicting comparisons between BRCA1 transcripts and isoforms by using SnapGene viewer software. Open reading frames (ORFs; the nucleotide sequences that begin with a start codon and end at a stop codon with no stop codon, and are able to be translated) that encode all the BRCA1 isoforms were obtained via ORFfinder on NCBI. The three-dimensional (3-D) structure of the isoforms was detected via Phyre2 software. Using PSORTII and SOPMA tools, the subcellular localization and the secondary structures of the isoforms were studied, respectively. Concerning the physicochemical properties of the BRCA1 proteins, they were determined by ProtParam software. Finally, the phosphorylation status of the isoforms was detected via NetPhos2.0 tool.

Results

Expression of BRCA1 in breast cancer cell lines

The findings showed high mRNA levels for BRCA1 in breast cancer cell lines: MCF-7, T47D, MDA-MB-231 and MDA-MB-468 compared to the normal breast tissue. Among the cancer cell lines, the expression of T47D was the highest (Figure 1).

Bioinformatics analysis of BRCA1
BRCA1 gene and its mRNAs and proteins sizes

According to Ensembl software, the BRCA1 gene (ENSG00000012048) was 145.95 kb in size. There were 33 transcripts and 31 isoforms of BRCA1. The gene had 146 orthologues and 95 phenotypes. According to the GenBank tool, which belongs to NCBI software, the BRCA1 gene (U14680.1) had five major transcripts with the sixth one as a non-protein coding, along with five isoforms. Among the principal transcripts, the largest one was 7287 bp long and the smallest was 3699 bp. Concerning the main isoforms, the longest one had 1884 amino acids and the shortest of 699 amino acids. SnapGene viewer tool depicted the differences among the major variants (Figure 2).

ORF analysis

Open reading frame, which is the nucleotide sequence that starts from the start codon and ends with the stop codon, encodes a polypeptide chain and contains no stop codon that interrupts its translation. As there were more than one coding DNA sequence (CDS) of the BRCA1 gene (31 CDSs according to Ensembl software), numerous ORFs were demonstrated using ORFfinder on NCBI. The analysis showed the presence of 46 ORFs for the transcript variant 1, with the +2 ORF was the longest with 5592 bp long and encodes 1863 amino acids. While for the transcript 2, there were also 46 ORFs, with the longest was +2 ORF of 5655 bp long, which encodes 1884 amino acids. Concerning transcript 3, there existed 47 ORFs, the longest of which was +3 ORF with the length of 5451 bp and encoded an amino acid of 1816 long. The transcript 4 demonstrated 27 ORFs with +2 ORF was the longest of 2280 bp encoding 759 amino acids. Regarding transcript 5, there were 26 ORFs, the longest of which was +3 ORF with the length of 2100 bp and encoded 699 amino acids.
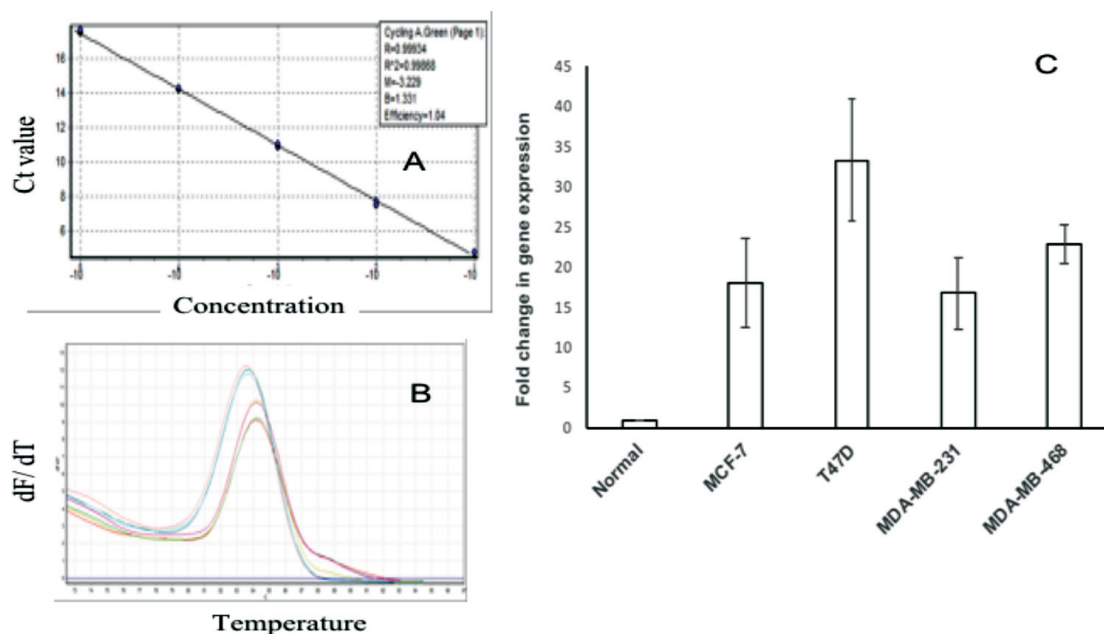

Figure 1. BRCA1 is overexpressed in cancer cell lines versus normal breast tissue. A- A standard curve plotted against Ct values of 18S and cDNA 10-fold serial dilutions of one cell line to calculate the amplification efficiency. B- An example of Melt curve without primer dimers but only 18S and BRCA1 cDNA expression in T47D and MDA-MB-231. C- Fold change in the BRCA1 mRNA expression in the breast carcinoma cells and normal breast tissue using RT-qPCR. Error bars represent the mean± standard deviation.

disordered.

Structure prediction of the isoforms
The secondary structure

The online software tool SOPMA predicted the secondary structure of BRCA1 encoding products. Table 1 and figure 3 below show details of the secondary structure of the the major BRCA1 isoforms.

The analytic results of Phyre2 software indicated that the secondary structure of BRCA1 isoforms was unique and did not resemble any other proteins except for short sequences that represented BRCT repeats of BRCA1 bound to a CTIP2 phosphopeptide. With regard to isoform 1, 11% of the sequence was similar to BRCT repeats of BRCA1 with 80% of the sequence was disordered. Similarly, 11% of the isoform 2 sequence was identical to the structure of the BRCT repeats, and 77% of the sequence was disordered. The sequence of isoform 3 showed 12% similarity with the structure of BRCT repetition of BRCA1 with a prediction of disordered sequence in 85% of it. While 28% of the isoform 4 sequence resembled the structure of the BRCT repeats and 52% of the sequence was predicted to be disordered, 16% of the sequence of isoform 5 was found to be similar to BRCT domain and 55% of the sequence was

The 3-D structure

Figure 4 shows the 3-D structures of the main BRCA1 spice variants analyzed by Phyre2 tool.

The subcellular localization

The PSORTII tool analyzed the subcellular localization of the BRCA1 proteins as shown in table 2.

Physicochemical characteristics

Some of the important physicochemical properties of BRCA1 proteins including relative molecular weight, theoretical isoelectric point (pI), the amino acid composition (Figure 5) and the instability index (II) were determined via ExPASy online software ProtParam as described below.

Isoform 1. The number of amino acids was 1863 with the molecular weight of 207.72 kDa and a pI of 5.29. The most prevalent amino acids were Ser (S) of 224 (12.0%) and Glu (E) of 198 (10.6%). Total number of positively charged residues (Asp+ Glu) was 283. Whereas the total number of the negatively charged residues (Arg + Lys) was 213. The II index was computed to be 54.68, which classified the protein as unstable.
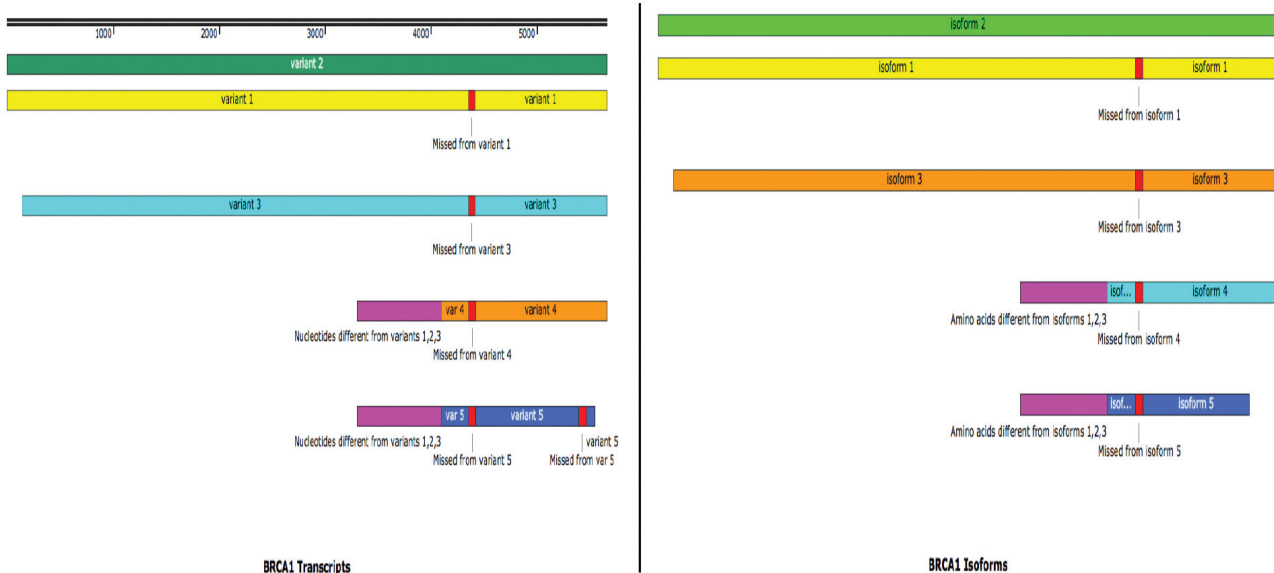
Isoform 2. It consisted of 1884 amino acids

Figure 2. Schematic diagrams show length and structure comparisons among BRCA1 transcripts (left) and splice variants (right). Red squares denote missing sequences, while pink areas refer to different regions.

Table 1. The secondary structure details of the BRCA1 isoforms

Isoform	α-helix	Extended strands	β-turns	Random coils
1	620 (33.28%)	189 (10.14%)	51 (2.74%)	1003 (53.84%)
2	630 (33.44%)	202 (10.72%)	49 (2.60%)	1003 (53.24%)
3	602 (33.15%)	187 (10.30%)	52 (2.86%)	975 (53.69%)
4	292 (38.47%)	95 (12.52%)	33 (4.35%)	339 (44.66%)
5	252 (36.05%)	84 (12.02%)	31 (4.43%)	332 (47.50%)

with the most were Ser (S) of 228 (12.1%) and Glu (E) of 199 (10.6%). The total residue number (Arg+Lys) of the positive charge was 217, whereas the total residue number (Asp+ Glu) of the negative charge was 285.This protein had a molecular weight of 210.27 kDa, pI of 5.36 and II of 55.02, which grouped the protein as unstable.

Isoform 3. It contained 1816 amino acids, the most common of which were Ser (S) counted 222 (12.2%), and Glu (E) numbered 193 (10.6%). The protein had a molecular weight of 202.34 kDa and pI of 5.30. Instability index (II) was found to be 54.38, which classified the protein as unstable. The total number of the negatively charged residues (Asp+ Glu) was 276, and the total number of the positively charged residues (Arg+ Lys) was 208.

Isoform 4. It had a molecular weight of 85.0 kDa and pI of 4.99. The amino acid composition of the 759 amino acid protein was mainly composed of 84 Ser (S), (11.1%), and 76 Glu (E) (10.0%). Although the total residue number (Arg+

Lys) of the positive charge was 76, the total residue number (Asp+ Glu) of the negative charge was 114. The II value was 54.4, which made the protein unstable.

Isoform 5. This protein possessed 699 amino acids, a molecular weight of 78.17 kDa and pI of 5.10. Similar to the other BRCA1 splice variants, the popular amino acids were Ser (S) and Glu (E) containing 79 (11.3%), and 70 (10.0%), respectively. The negatively charged residues (Asp+ Glu) accounted for 105, while the positively charged residues (Arg+ Lys) were 76. The value of II was 57.35, which offered the protein the instability.

Phosphorylation sites analysis

The phosphorylation sites of the BRCA1 isoforms were determined via NetPhos2.0 tool, which predicted the locations of the amino acids serine, threonine, and tyrosine of each protein. Each of isoform 1 and isoform 2 contained 188 serine, 67 threonine and 16 tyrosine. There were
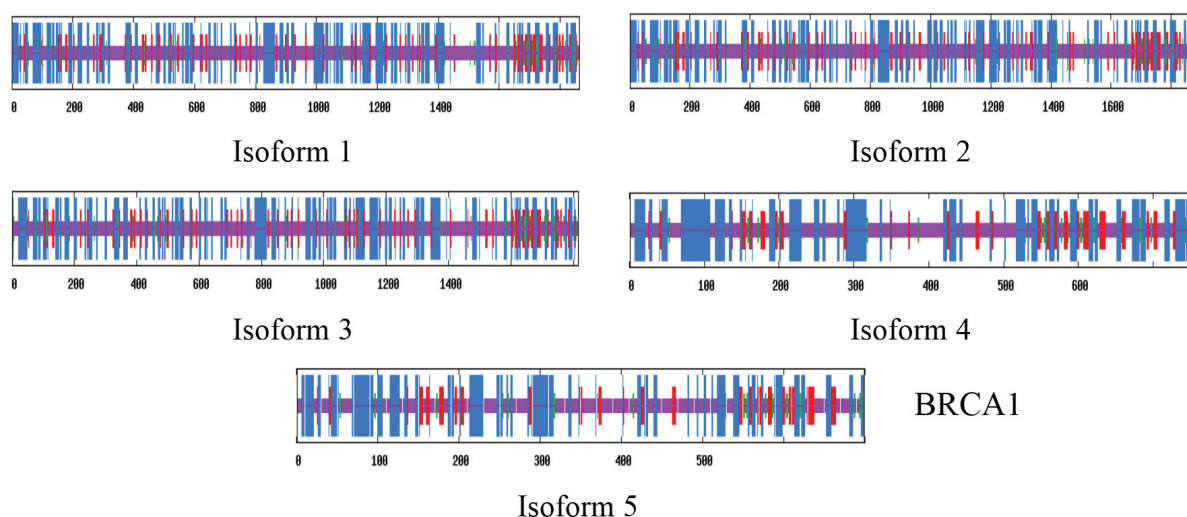

Figure 3. Secondary structure prediction of BRCA1 encoded products using SOPMA.

Table 2. Subcellular localization predicted for the BRCA1 isoforms

Subcellular localization	BRCA1 isoform (possibility %)				
	1	2	3	4	5
nuclear	87.0%	82.6%	87.0%	65.2%	65.2%
Cytoplasmic	4.3%	8.7%	4.3%	21.7%	21.7%
Plasma membrane	4.3%	4.3%	4.3%	4.3%	4.3%
Vesicles of secretory system	4.3%	4.3%	4.3%	4.3%	4.3%
Mitochondrial	-	-	-	4.3%	4.3%

182 serine, 66 threonine 16 tyrosine in isoform 3, 69 serine , 33 threonine, and 10 tyrosine in isoform 4. Also, 64 serine, 30 threonine and 9 tyrosine in isoform 5.

Phylogenetic tree

For the phylogenetic analysis, the BRCA1 coding products of different species that showed 100 up to 95% identity with homo sapiens BRCA1 were collected via BLAST search from NCBI software. The evolutionary relationships of BRCA1 were inferred using the UPGMA method. The tree was drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The analysis involved 11 amino acid sequences (from 11 organisms). All positions containing gaps and missing data were eliminated. There were a total of 1860 positions in the final dataset. The evolutionary distances between BRCA1 of human and Chimpanzee (Pan troglodytes) was the shortest and the closest to the genetic relationship. Evolutionary analyses were conducted by MEGA7 (Figure 6).

Discussion

The data of this study show the overexpression of the BRCA1 mRNA in the four studied familial breast cancer cell lines including: MCF7 and T47D (both are estrogen receptor ER positive,
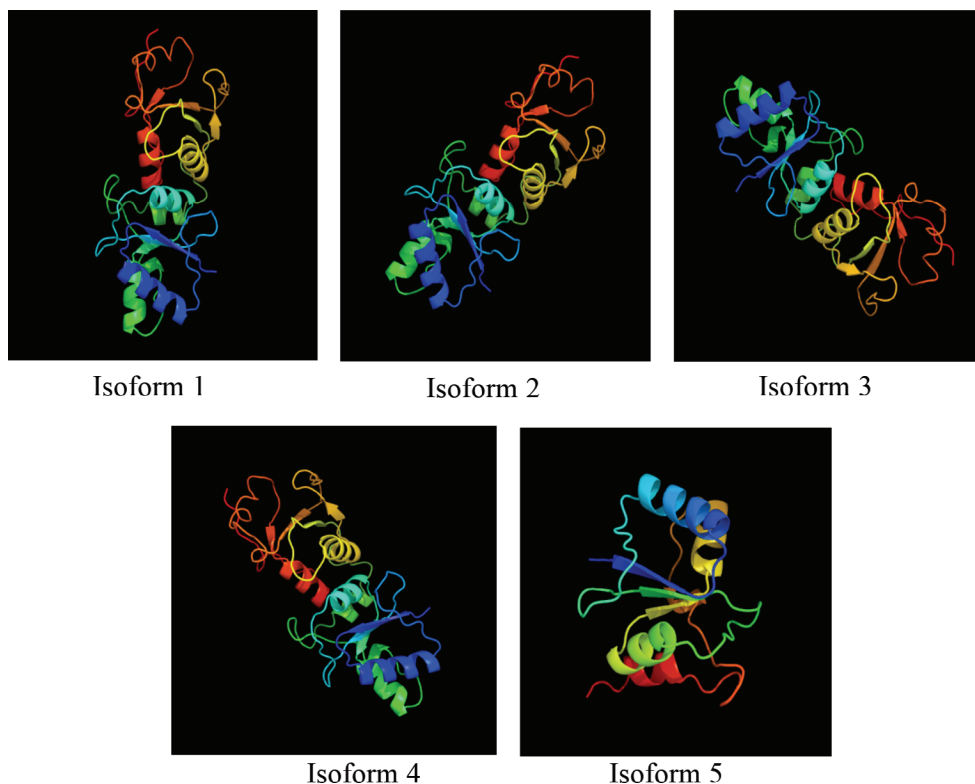

Isoform 1   Isoform 2   Isoform 3

Isoform 4   Isoform 5

Figure 4. Three-dimensional structures of the BRCA1 splice variants.

progesterone receptor PR positive and human epidermal growth factor receptor HER2 negative) as well as MDA-MB-468 and MDA-MB-231 (both are triple negative) compared to the lower expression of the normal breast tissue cDNA. Partially, these results are similar to those of Gudas et al.[29] who found BRCA1 transcripts to be highly expressed in MCF-7, T47D and MDA-MB-468, but not in MDA-MB-231 cells whose mRNA levels were relatively low.[29,30] It has been suggested that improper expression of BRCA1, instead of somatic mutation, possibly participate in some sporadic cases of human breast neoplasms.[31] Notably, sporadic breast tumor tissues have not reported somatic BRCA1 mutations.[32,33] As a tumor suppressor, BRCA1 was thought to participate in carcinogenesis by

decreasing in its mRNA transcription and protein translation along with alterations of promoter methylation.[34-36] However, BRCA1 overexpression registered in this study might be functionally impaired, and the increased BRCA1 transcripts levels have been proposed as an attempt of the cells to compensate for their loss of BRCA1 functions. This hypothesis was supported by the findings of MCF-7 cells, whose endogenous BRCA1 expression despite showing elevated steady-state BRCA1 endogenous levels.[37] From another point of view, as BRCA1 gene is located on chromosome 17, which has frequently been found to be amplified in human breast malignancies, allelic losses of this gene at 17q are seldom estimated because techniques, depending on karyotype detection, fail to find
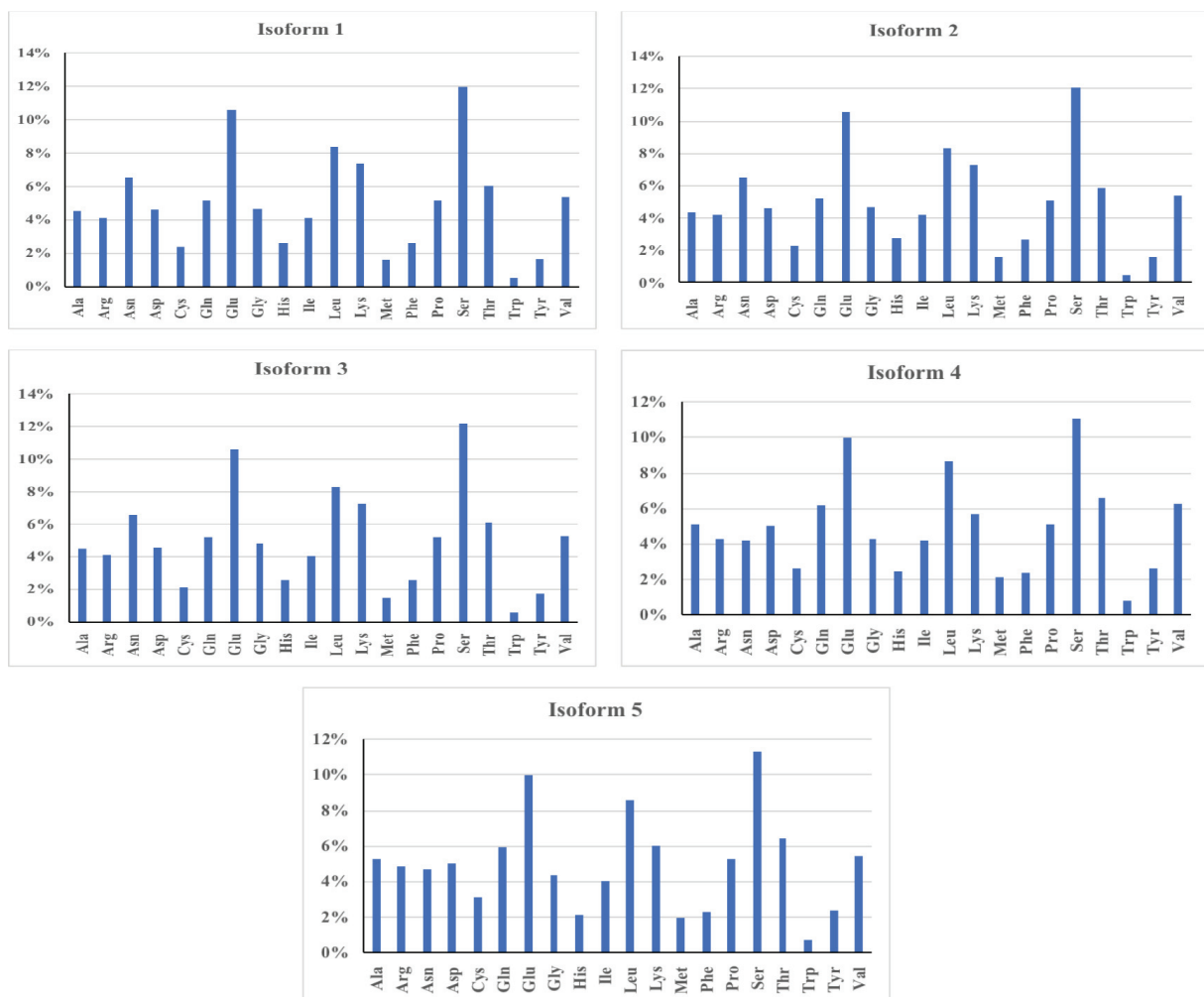

Figure 5. The total amino acids composition of the major BRCA1 isoforms.

loss, especially when the retained allele is re-duplicated or amplified.[38]

The high expression of BRCA1 transcripts is compatible with the findings of Elstrodt et al.[5] who analyzed BRCA1 mutations in 41 breast cancer cell lines. The cells used in our study possess the wild type BRCA1 gene, and its transcripts are unmethylated. Real time quantitative PCR applied in this study revealed the highest mRNA expression in T47D cells. This might be due to no allelic loss of this gene in the aforementioned cell line when compared with BRCA1 allele loss in the other three breast cancer cell types.[5] Generally, high expression of the BRCA1 mRNA has been demonstrated to be correlated with the increased proliferation.[29] Similar to BRCA1 overexpression, the other BRCA1 product, known as IRIS, has been highly expressed in multiple human cancer cell lines. Hence, IRIS can act as a tumor-progression driver in many cancers when spontaneously overexpressed.[2]

Bioinformatics analysis of BRCA1 transcripts and splice variants was investigated in this study to shed light on many of their features that might help the laboratory research and provide explanations to some controversial experimental findings. Interestingly, different species of BRCA1 mRNAs and proteins have been proposed to play different biological roles, and their essential normal functions could be maintained by keeping their appropriate ratio.[29] The bioinformatics analysis of this study shows that 33 BRCA1 transcripts and 31 isoforms can be found in human being, of which there are five major mRNAs and proteins extensively analyzed here. According to NCBI, the largest transcript is 7287 bp long (transcript variant 2) and the smallest is 3699 bp long (transcript 4), while the longest isoform is isoform 2 whose length is 1884 amino acids and the shortest is isoform 5 (699 amino acids). The presence of many transcript and splice variants of this gene occurs because BRCA1 has multiple ORFs. The secondary structure of the BRCA1 main isoforms was predicted via bioinformatics tools exemplified by SOPMA, illustrated the percentages of each of α-helix, β-turns, extended strands and random coils of the major products encoded by BRCA1. Interestingly, random coils constitute about 50% of the structure of these isoforms followed by α-helix (around 35%). Also, the tool used to study BRCA1 secondary structure was called Phyre2, which demonstrated unique structures that are not similar to other products. Except for short regions located in all of the five major isoforms that were shown to resemble BRCT repetitions of BRCA1 bound to a CTIP2 phosphopeptide. In this regard, BRCA1 has been
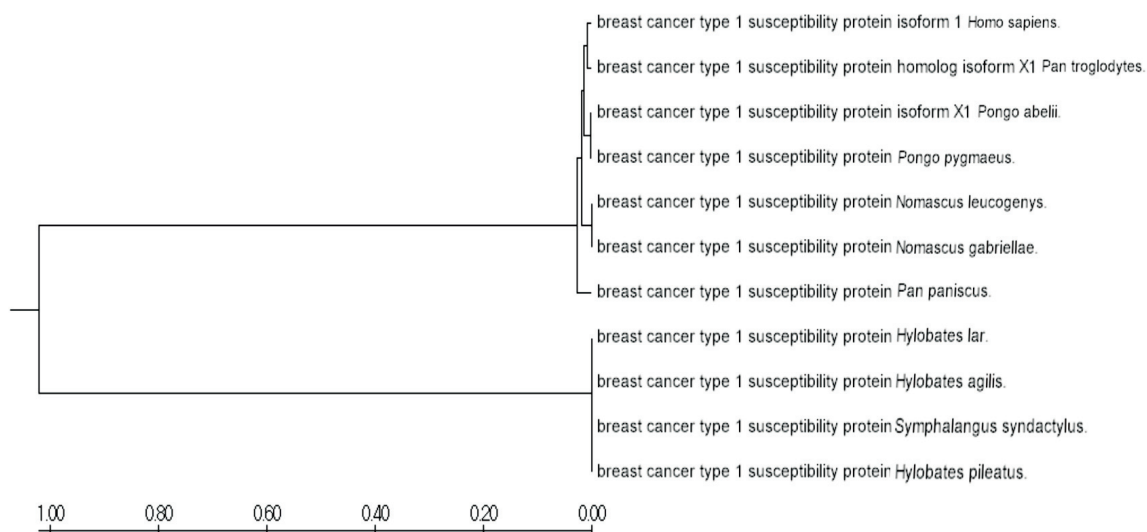

Figure 6. Phylogenetic tree of the BRCA1 encoding products of various species. The optimal tree with the sum of branch length = 2.11645655 is shown.

designated to have several functional domains, including one N-terminal RING finger domain, two nuclear localization signal domains, and two BRCA1 carboxy terminal domains referred to as BRCT motifs,[39,40] functioning as a strong transcriptional activator when fused to a heterologous DNA binding domain.[41] Most importantly, the N-terminal domain was not in the variants 4 and 5, and parts of the C-terminal domain were missing from these variants (Figure 2). Truncation or loss of the C-terminal BRCT domain has been associated with almost all BRCA1 mutations in germ-line cells, which suggests an important function of the gene in transcriptional regulation.[23]

Regarding the subcellular localization, a bioinformatics tool showed nuclear persistence of the BRCA1 proteins. Although isoforms 4 and 5 also revealed 21.7% cytoplasmic localization, the lower percentages of 8.7% and 4.3% were shown by variants 2 and 1, respectively. In addition, BRCA1 can also be found in other cell compartments, such as plasma membrane and vesicles of secretory system but in low percentages. Only can variants 4 and 5 be seen in mitochondria. Chen et al. observed the unusual subcellular localization of BRCA1 protein in breast cancer cells.[31] This protein showed spotted nuclear pattern in many breast cancer cell lines, centrosomes, and mitochondria in immunohistochemistry assay by using MS110 antibody.[42] Similarly, four monoclonal anti-BRCA1 antibodies stained the nucleus of breast cancer cell lines.[43] Concerning frozen and formalin-fixed paraffin-embedded (FFPE) tissue sections of breast tumors tested by immunohistochemistry, they showed nuclear and nucleolar localization of BRCA1 protein by using the monoclonal antibodies MS110 and AP16.[1] However, FFPE specimens made from benign breast, invasive lobular cancers, and low grade ductal cancer revealed identical nuclear BRCA1 staining, whereas breast cancer tissue from patients with mutated BRCA1 as well as patients with high-grade ductal carcinomas had almost no staining.[44] Furthermore, BRCA1 was found mainly in the cytoplasm of malignant mammary cells, and detected in the nuclei of epithelial cells.[31]

Conversely, BRCA1 was shown to be located mostly in the nuclei of normal and cancer cells as well.[43]

Thus, the staining of BRCA1 protein has been the subject of debate, and these controversial staining pattern might be due to differences in staining procedures, variations in antibodies specificities adding to staining of lymphocytes in surrounding stromal tissue.[44] Zhang et al. added that the BRCA1 subcellular localization can be frequently impeded by technical errors ascribed to loss specificity of certain BRCA1 antibodies and cross-reactivity with other epitopes. With respect to antibody specificity variations, this situation possibly can be solved by the use of bioinformatics tools to analyze BRCA1 isoforms, this might be helpful in choosing appropriate antibodies that target epitopes common to all the variants or even specific to the certain isoform if the author is interested to know the cellular distribution of that BRCA1 protein.

Importantly, among 19 anti-BRCA1 antibodies, the monoclonal MS110 (Ab-1) antibody targeting the first 304 amino acids of BRCA1 at its N-terminus is suggested to be the highly specific antibody to detect the BRCA1 protein levels and distribution in normal and neoplastic human breast and ovarian tissues.[46] However, recently, with the advance of molecular technologies, BRCA1 has been demonstrated to be a shuttle protein, transported between specific sites within the nucleus and cytoplasm involving centrosomes, mitochondria and DNA repair foci[42] although little is known about the regulation of its shuttling between the nucleus and cytoplasm.[47]

The physicochemical properties of BRCA1 encoding products were investigated via bioinformatics tools. The five major isoforms varied in their molecular weights from 210 kDa of the longest isoform to 78 kDa of the shortest. Although, the molecular weight of isoforms 1 and 2 is of approximate value of 207.7 kDa and 210 kDa, respectively, the most commonly reported BRCA1 protein is of 220 kDa. It has been found that BRCA1 proteins have different molecular weights from those expected from the amino acids sequences encoded by BRCA1

mRNA.[29] Regardless of the isoform, the most common amino acids are Ser and Glu and all of the isoforms are unstable due to their increased instability index.

It is well-known that BRCA1 is a serine phosphoprotein regulated in response to DNA damage.[48] Serine in position 1423 has been identified to be phosphorylated among several phosphorylation sites via phospho-Ser-specific antibodies.[45] Using NetPhos2.0 tool, we were able to determine a large number of predicted phosphorylation sites in BRCA1 splice variants, with isoforms 1 and 2 contained larger numbers of serine that are predicted to be phosphorylated from isoforms 3, 4, and 5. Increased nuclear phosphorylated BRCA1 was shown to be expressed by more than 70% of triple negative breast cancer cells, BRCA1-mulated or deficient cells, although it is inefficient protein.[49]

In the last part of this study, the BRCA1 phylogenetic relationship was analyzed using a special bioinformatics program. Among the 11 amino acid sequences aligned that represented different primates and organisms, human and chimpanzee BRCA1 sequences were very close indicating genetic relatedness. In this respect, chimpanzees can be used as a model to design therapies that might help curing humans. To conclude, BRCA1 overexpression in the breast carcinoma cell lines might be very essential in the pathogenesis of sporadic breast cancer. Bioinformatics approaches applied in this study, can be used as alternative or supportive methods to the laboratory molecular techniques. The bioinformatics analysis of BRCA1 mRNA and protein variants can provide vital information for cancer diagnosis or therapy. Moreover, the biological data gained from the bioinformatics tools can help authors make better decisions before launching expensive experiments.

Conflict of Interest

None declared.

References

1. Tulchin N, Ornstein L, Dikman S, Strauchen J, Jaffer S, Nagi C, et al. Localization of BRCA1 protein in breast cancer tissue and cell lines with mutations. *Cancer Cell Int*. 2013;13(1):70. doi: 10.1186/1475-2867-13-70.
2. Li AG, Murphy EC, Culhane AC, Powell E, Wang H, Bronson RT, et al. BRCA1-IRIS promotes human tumor progression through PTEN blockade and HIF-1α activation. *Proc Natl Acad Sci USA*. 2018;115(41):E9600-E9609. doi: 10.1073/pnas.1807112115.
3. Tammaro C, Raponi M, Wilson DI, Baralle D. BRCA1 exon 11 alternative splicing, multiple functions and the association with cancer. *Biochem Soc Trans*. 2012;40(4):768-72.
4. Hill SJ, Clark AP, Silver DP, Livingston DM. BRCA1 pathway function in basal-like breast cancer cells. *Mol Cell Biol*. 2014;34(20):3828-42.
5. Elstrodt F, Hollestelle A, Nagel JH, Gorin M, Wasielewski M, van den Ouweland A, et al. BRCA1 mutation analysis of 41 human breast cancer cell lines reveals three new deleterious mutants. *Cancer Res*. 2006;66(1):41-5.
6. Zheng L, Pan H, Li S, Flesken-Nikitin A, Chen PL, Boyer TG, et al. Sequence-specific transcriptional corepressor function for BRCA1 through a novel zinc finger protein, ZBRK1. *Mol Cell*. 2000;6(4):757-68.
7. Xie A, Kwok A, Scully R. Role of mammalian Mre11 in classical and alternative nonhomologous end joining. *Nat Struct Mol Biol.* 2009;16(8):814-8.
8. Venkitaraman AR. Cancer susceptibility and the functions of BRCA1 and BRCA2. *Cell.* 2002;108(2):171-82.
9. Tulchin N, Chambon M, Juan G, Dikman S, Strauchen J, Ornstein L, et al. BRCA1 protein and nucleolin colocalize in breast carcinoma tissue and cancer cell lines. *Am J Path.* 2010;176:1203-14.
10. Guerra-Rebollo M, Mateo F, Franke K, Huen MS, Lopitz-Otsoa F, Rodriguez MS, et al. Nucleolar exit of RNF8 and BRCA1 in response to DNA damage. *Exp Cell Res.* 2012;318(18):2365-76.
11. Walsh T, King MC. Ten genes for inherited breast cancer. *Cancer Cell.* 2007;11(2):103-5.
12. Li ML, Greenberg RA. Links between genome integrity and BRCA1 tumor suppression. *Trends Biochem Sci.* 2012;37(10):418-24.
13. Silver DP, Livingston DM. Mechanisms of BRCA1 tumor suppression. *Cancer Discov.* 2012;2:679-84.
14. Pathania S, Nguyen J, Hill SJ, Scully R, Adelmant GO, Marto JA, et al. BRCA1 is required for postreplication repair after UV-induced DNA damage. *Mol Cell.* 2011;44(2):235-51.
15. Schlacher K, Wu H, Jasin M. A distinct replication fork protection pathway connects Fanconi anemia tumor suppressors to RAD51-BRCA1/2. *Cancer Cell.* 2012;22(1):106-16.
16. Garcia-Higuera I, Taniguchi T, Ganesan S, Meyn MS, Timmers C, Hejna J, et al. Interaction of the Fanconi

anemia proteins and BRCA1 in a common pathway. *Mol Cell.* 2001;7:249-62.

17. Vandenberg CJ, Gergely F, Ong CY, Pace P, Mallery DL, Hiom K, et al. BRCA1-independent ubiquitination of FANCD2. *Mol Cell.* 2003;12(1):247-54.

18. Bunting SF, Callen E, Kozak ML, Kim JM, Wong N, Lopez-Contreras AJ, et al. BRCA1 functions independently of homologous recombination in DNA interstrand crosslink repair. *Mol Cell.* 2012;46(2):125-35.

19. Joukov V, Groen AC, Prokhorova T, Gerson R, White E, Rodriguez A, et al. The BRCA1/BARD1 heterodimer modulates ran-dependent mitotic spindle assembly. *Cell.* 2006;127(3):539-52.

20. Xu X, Weaver Z, Linke SP, Li C, Gotay J, Wang XW, et al. Centrosome amplification and a defective G2-M cell cycle checkpoint induce genetic instability in BRCA1 exon 11 isoform-deficient cells. *Mol Cell.* 1999;3(3):389-95.

21. Hsu LC, Doan TP, White RL. Identification of a gamma-tubulin-binding domain in BRCA1. *Cancer Res.* 2001;61(21):7713-8.

22. Zhu Q, Pao GM, Huynh AM, Suh H, Tonnu N, Petra M, et al. BRCA1 tumour suppression occurs via heterochromatin-mediated silencing. *Nature.* 2011;477(7363):179-84.

23. Atalay A, Crook T, Ozturk M, Yulug IG. Identification of genes induced by BRCA1 in breast cancer cells. *Biochem Biophys Res Commun.* 2002;299(5):839-46.

24. ElShamy WM, Livingston DM. Identification of BRCA1-IRIS, a BRCA1 locus product. *Nat Cell Biol.* 2004;6(10):954-67.

25. Nakuci E, Mahner S, Direnzo J, ElShamy WM. BRCA1-IRIS regulates cyclin D1 expression in breast cancer cells. *Exp Cell Res.* 2006;312(16):3120-31.

26. Mehmood MA, Sehar U, Ahmad N. Use of bioinformatics tools in different spheres of life sciences. *J Data Mining in Genomics & Proteomics.* 2014;5(2).

27. Bayat A. Science, medicine, and the future: Bioinformatics. *BMJ.* 2002;324(7344):1018-22.

28. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. *Methods.* 2001;25(4):402-8.

29. Gudas JM, Li T, Nguyen H, Jensen D, Rauscher FJ, Cowan KH. Cell cycle regulation of BRCA1 messenger RNA in human breast epithelial cells. *Cell Growth Differ.* 1996;7(6):717-23.

30. Blagosklonny MV, An WG, Melillo G, Nguyen P, Trepel JB, Neckers LM. Regulation of BRCA1 by protein degradation. *Oncogene.* 1999;18(47):6460-8.

31. Chen Y, Chen CF, Riley DJ, Allred DC, Chen PL, Von Hoff D, et al. Aberrant subcellular localization of BRCA1 in breast cancer. *Science.* 1995;270(5237):789-91.

32. Lindblom A, Skoog L, Rotstein S, Werelius B, Larsson C, Nordenskjold M. Loss of heterozygosity in familial breast carcinomas. *Cancer Res.* 1993;53(18):4356-61.

33. Futreal PA, Liu Q, Shattuck-Eidens D, Cochran C, Harshman K, Tavtigian S, et al. BRCA1 mutations in primary breast and ovarian carcinomas. *Science.* 1994;266(5182):120-2.

34. Rio PG, Maurizis JC, Peffault de Latour M, Bignon YJ, Bernard-Gallon DJ. Quantification of BRCA1 protein in sporadic breast carcinoma with or without loss of heterozygosity of the BRCA1 gene. *Int J Cancer.* 1999;80(6):823-6.

35. Egawa C, Miyoshi Y, Taguchi T, Tamaki Y, Noguchi S. Quantitative analysis of BRCA1 and BRCA2 mRNA expression in sporadic breast carcinomas and its relationship with clinicopathological characteristics. *Jpn J Cancer Res.* 2001;92(6):624-30.

36. Guendel I, Carpio L, Pedati C, Schwartz A, Teal C, Kashanchi F, et al. Methylation of the tumor suppressor protein, BRCA1, influences its transcriptional cofactor function. *PLoS One.* 2010;5(6):e11379.

37. Holt JT, Thompson ME, Szabo C, Robinson-Benion C, Arteaga CL, King MC, et al. Growth retardation and tumour inhibition by BRCA1. *Nat Genet.* 1996;12(3):298-302.

38. Harkes IC, Elstrodt F, Dinjens WN, Molier M, Klijn JG, Berns EM, et al. Allelotype of 28 human breast cancer cell lines and xenografts. *Br J Cancer.* 2003;89(12):2289-92.

39. Scully R, Chen J, Ochs RL, Keegan K, Hoekstra M, Feunteun J, et al. Dynamic changes of BRCA1 subnuclear location and phosphorylation state are initiated by DNA damage. *Cell.* 1997;90(3):425-35.

40. Wang Y, Cortez D, Yazdi P, Neff N, Elledge SJ, Qin J. BASC, a super complex of BRCA1-associated proteins involved in the recognition and repair of aberrant DNA structures. *Genes Dev.* 2000;14(8):927-39.

41. Monteiro AN, August A, Hanafusa H. Evidence for a transcriptional activation function of BRCA1 C-terminal region. *Proc Natl Acad Sci U S A.* 1996;93(24):13595-9.

42. Henderson BR. The BRCA1 breast cancer suppressor: Regulation of transport, dynamics, and function at multiple subcellular locations. *Scientifica (Cairo).* 2012;2012:796808.

43. Scully R, Ganesan S, Brown M, De Caprio JA, Cannistra SA, Feunteun J, et al. Location of BRCA1 in human breast and ovarian cancer cells. *Science.* 1996;272(5258):123-6.

44. Rakha EA, El-Sheikh SE, Kandil MA, El-Sayed ME, Green AR, Ellis IQ. Expression of BRCA1 protein in breast cancer and its prognostic significance. *Hum Pathol.* 2008;39(6):857-65.

45. Zhang X, Hofmann S, Harbeck N, Jeschke U, Sixou S. Impact of etoposide on BRCA1 expression in various breast cancer cell lines. *Drugs R D.* 2017;17(4):569-

    83.

46. Wilson CA, Ramos L, Villasenor MR, Anders KH, Press MF, Clarke K, et al. Localization of human BRCA1 and its loss in high-grade, non-inherited breast carcinomas. *Nat Genet.* 1999;21(2):236-40.

47. Feng Z, Kachnic L, Zhang J, Powell SN, Xia F. DNA damage induces p53-dependent BRCA1 nuclear export. *J Biol Chem.* 2004;279(27):28574-84.

48. Okada S, Ouchi T. Cell cycle differences in DNA damage-induced BRCA1 phosphorylation affect its subcellular localization. *J Biol Chem.* 2003;278(3): 2015-20.

49. Uhlen M. Response to: Should we ignore western blots when selecting antibodies for other applications? *Nat Methods.* 2017;14(3):215-6.