Acta Scientific CANCER BIOLOGY (ISSN: 2582-4473)

Volume 7 Issue 8 October 2023

Research Article

BRCA1/2-mutant Breast Cancer: Genome-Wide Comparative Analysis, Differentially Expressed Genes and Key Molecules Analyze by Evidence from Bioinformatics Analyses

Ghania Qindeel*

Government College University, Lahore, Pakistan

***Corresponding Author:** Ghania Qindeel, Government College University, Lahore, Pakistan.

DOI: 10.31080/ASCB.2023.07.0443

Received: August 17, 2023

Published: September 06, 2023

Abstract

The *BRCA1* and *BRCA2* genes have been linked to a higher risk of breast cancer throughout a person's lifespan. *BRCA1/2* mutations in breast cancer have yet to be thoroughly characterized in terms of the genes that are closely linked to those alterations. This project seeks to discover gene expression interaction and patterns networks affected by *BRCA1/2* mutations, to represent underlying disease processes and provide novel biomarkers for breast cancer diagnosis or prognosis. The advancement of molecular genetics in recent years has improved our understanding of the concepts of breast cancer development. This study is accomplished to understand genomic variations and for sequence analysis. Freely accessible online tools such as pfam, GSDS (gene structure display server), SMART, STRING 9.1, PTMcode 2,PDB, MEME motif, CLUSTAL W, Serial Cloner and Mega X were used to perform sequence and mutational analysis. The Cancer Genome Atlas (TCGA), gene expression profiles were obtained and integrated with the cBioPortal website to identify specific breast cancer patients with *BRCA1/2* mutations, according to the study. Using gene set enrichment analysis (GSEA), certain enriched pathways and biochemical characteristics linked with BRCA mutations were identified and characterized. Three separate differentially expressed genes (DEGs) analyses were done for *BRCA1/2*-mutant breast cancer, wild-type breast cancer, and the matching normal tissues to validate putative hub genes with each other. Key genes linked with *BRCA1/2* mutations were identified using protein-protein interaction networks (PPI), survival analysis, and diagnostic value evaluation. Taken together, our findings shed light on specific mutations and proteins implicated in the interaction network of *BRCA1* and *BRCA2*, which may play similar roles in breast cancer. However, the complicated process behind these results has yet to be fully explained, and further research is needed in the future.

Keywords: Mutations; BRCA; Biomarkers; Bioinformatics

Introduction

The *BRCA1* and *BRCA2* genes have been generally recognized as the major penetrant genes that cause a genetic susceptibility to breast cancer syndrome since their identification in the 1990s [1-3]. Breast cancer risk is expected to be 57 percent to 65 percent and 45 percent to 55 percent in women with *BRCA1* and *BRCA2* mutations, respectively, whereas the ovarian cancer risk is estimated to be 39 percent to 44 percent and 11 percent to 18 percent in women with *BRCA1* and *BRCA2* mutations [4-6]. *BRCA1* and *BRCA2* are tumor suppressor genes. *BRCA1* is a complex gene having 24 exons found on chromosome 17 (17q21) that codes for a 1,863 amino acid zinc-binding protein. Gene finger motif at the amino terminus and a conserved acidic carboxyl-terminal motif (*BRCA1* C terminal) make this a really interesting gene [1,7,8]. On

chromosome 13 (13q12), the *BRCA2* gene that codes for a 3,418 amino acid protein that is structurally and functionally identical to the *BRCA1*. Researchers believe that *BRCA1* and *BRCA2* proteins have a role in double-strand DNA repair via sustaining genomic integrity via RAD51 [9] and that they are also actively involved in the homologous recombination process [10].

There have been almost 165 mutations and sequence variations identified in the *BRCA1* gene, and 1730 in the *BRCA2* gene [11]. Cancer-suppressing *BRCA1* and *BRCA2* genes both have a role in DNA repair. The *BRCA1* gene is also involved in the regulation of checkpoints. Autosomal dominant breast cancer is related to mutations in either gene [13].

In order to reflect the pathogenicity of mutations in these domains, studies have assessed the specificity and sensitivity of functional assays for the *BRCA1* BRCT domain and the *BRCA2* DNA-binding domain [14-18]. This research contrasted functional assay findings against pathogenicity given only based on clinical data, and they created the foundation for incorporating data from these particular functional assays into the multifactorial probability model. Other domains of *BRCA1* and *BRCA2* are known to be crucial for function, such as the RING, transcriptional activation, and *BRCA1* c-terminal domains. Furthermore, there have been no systematic studies to date that have calibrated the degree of function of variations in these domains toward clinical information to assess the specificity and sensitivity of relevant tests to indirectly analyze cancer risk. Furthermore, the calibration and development of quantitative splicing assays toward direct risk measures will be critical for improving bio-informatic prediction tools, improving the estimation of bio-informatically defined prior probabilities, and allowing the acceptance of mRNA assay data as a likelihood component of the multi-factorial model. For 19 *BRCA1* and *BRCA2* exonic variants, we describe the findings of multifactor probability modeling and/or bio-informatically targeted splicing tests to give variant categorization of direct clinical value. They can be used as a calibration set of variants for future studies that update bio-informatically estimated prior probabilities of pathogenicity for variants and incorporate splicing and functional assays into multifactorial likelihood estimates.

Supplementary material

Proposed name	Chro-mo-some #	Gene Locus	Exons	ORF size	Amino acid length	Protein ac-cession #	RNA accession #	Start of genomic Location	Conserved domains in protein sequence
BRCA1	17	NC_000017.11	24	5589bp	1863	NP_009225.1	NM_007294.4	43014607	the N-terminal RING domain, exons 11-13, and the BRCT domain
BRCA2	13	NC_000013.11	27	10254bp	3418	NP_000050.3	NM_000059.4	32031774	protein containing domains BRCA-2 helical, BRCA2DBD_ OB1, BRCA2D-BD_OB2, and BRCA2DBD_ OB3

Table a

Materials and Methods

BRCA genes identification

The *BRCA1* and *BRCA2* genes' genomes and protein sequences were acquired from NCBI (www.ncbi.nlm.nih.gov). The protein sequences of the human *BRCA1* and *BRCA2* genes were used in the protein blast. Other species were used to align. The E-value of the variants with the longest open reading frames was carefully verified. Some online servers was used to remove sequences that did not include conserved motifs, such as the SMART (http://smart.embl-heidelberg.de/) database and the Pfam (http://pfam.janelia.org/) database. Sequences with conserved domains were identified by running them through some of these databases. The NCBI website was used to acquire information such as chromosomal position, start of genomic sequence (*BRCA1* and *BRCA2*), the number of amino acids, RNA and protein accession numbers, gene locus, and length of open reading frames.

Figure 1: Substantial matches of motifs of BRCA1 gene (a) and BRCA2 gene (b) from different species using Pfam.

Posttranslational modifications (PTMs) of BRCA genes

The *BRCA1* and *BRCA2* genes' protein sequences were acquired from NCBI (www.ncbi.nlm.nih.gov). The protein sequences of the human *BRCA1* and *BRCA2* genes were put into an online server, PTMcode 2 (https://ptmcode.embl.de/). In *BRCA1* protein, there are 50 PTMs annotated (4 propagated), with 398 functional associations. In *BRCA2* protein, there are 95 PTMs annotated (8 propagated), with 11 functional associations.

Identification of structure of gene

The structure of *BRCA 1* and *BRCA 2* gene was visualized by using GSDS (Gene structure Display Server; GSDS - http://gsds.gao-lab.org), a free web-based gene structure visualizing tool.

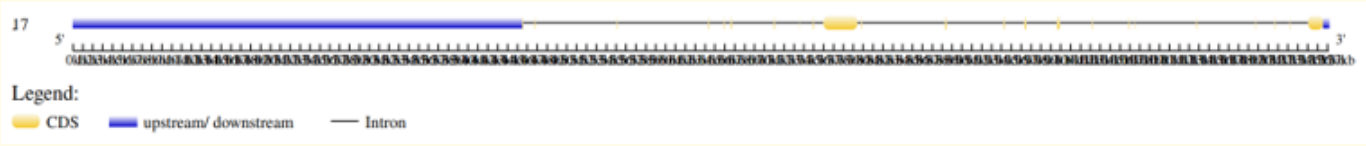

Figure 2: Gene structure of BRCA 1 (GSDS).

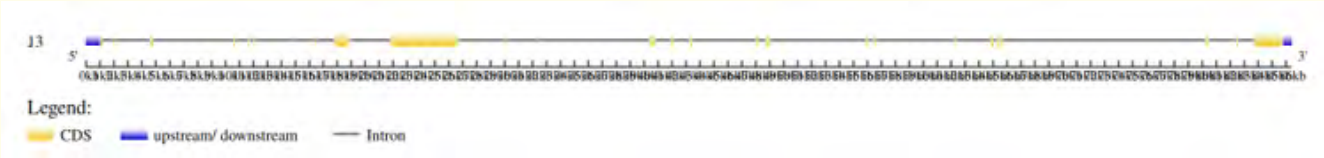

Figure 3: Gene structure of BRCA 2 (GSDS).

BRCA 1 gene is located on chromosome 17q21.31 and contains 24 exons; whilst, *BRCA 2* gene is located on Chromosome 13q13.1 and contains 27 exons. 3-D protein structure for major protein encoded by *BRCA 1* and *BRCA 2* gene was obtained from PDB (https://www.rcsb.org/).

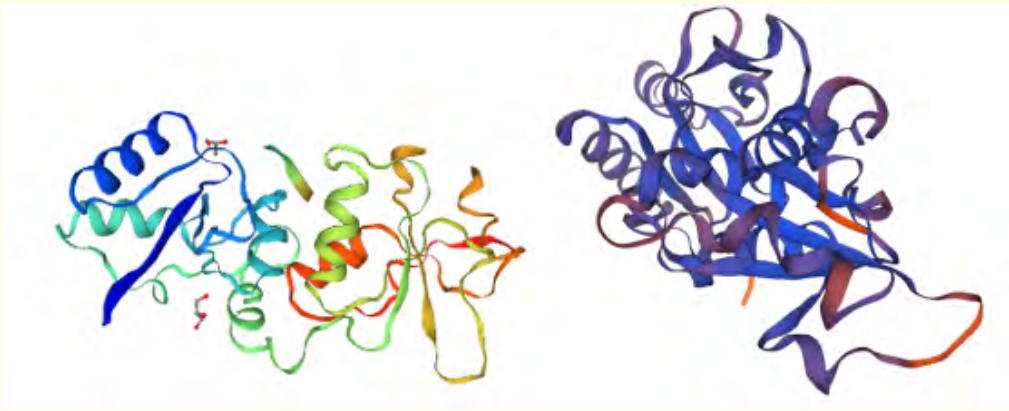

Figure 4: 3-D structure of major protein encoded by BRCA1 (left) and BRCA 2 (gene) in Humans (PDB)

2.4- Phylogenetic Analysis.

Phylogenetic analysis of both *BRCA 1* and *BRCA 2* gene was performed on MEGA X software (https://www.megasoftware.net/) as shown in figure 5 and 6. Neighbor-joining method was used for this experiment. Genomic sequences of *BRCA 1* and *BRCA 2* and their orthologs were obtained from NCBI (www.ncbi.nlm.nih.gov) and were subjected to Basic local alignment, most closely related sequences (having high E score) were selected for the construction of phylogenetic tree. 25 species were screened out for analysis of *BRCA 1* and *BRCA 2*, all sequences were obtained from NCBI.

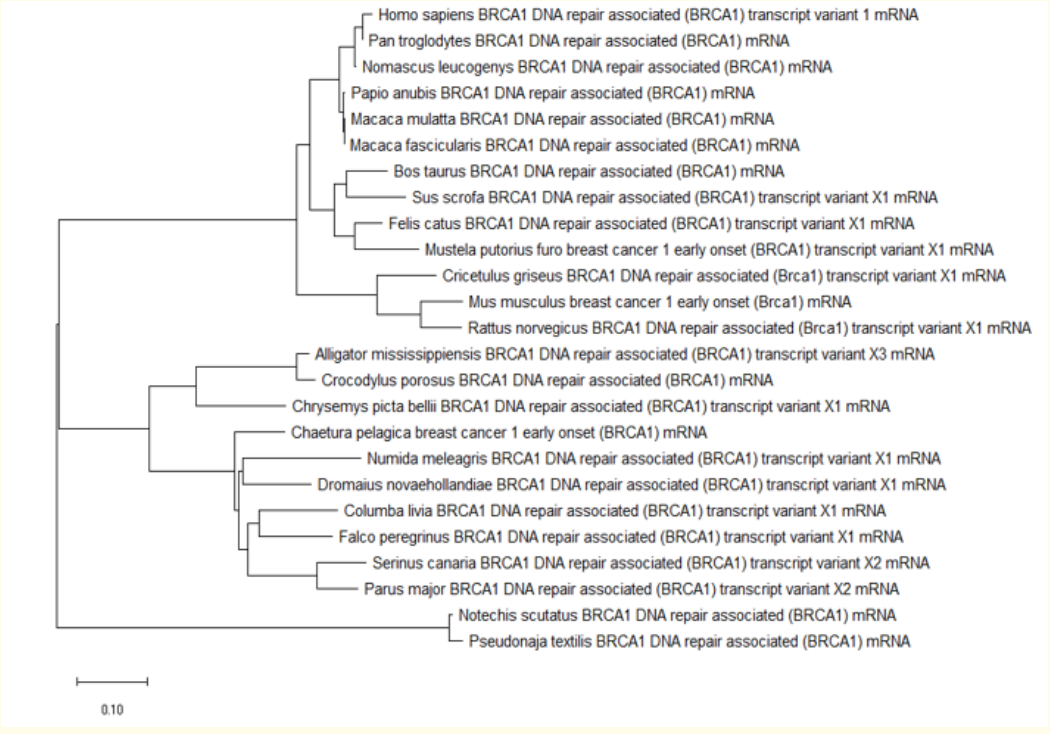

Figure 5: Phylogenetic analysis of BRCA 1 gene with its orthologs (MEGA X).

Figure 6: Phylogenetic analysis of BRCA 2 gene with its orthologs (MEGA X).

Identification of conserved motifs

Conserved motifs in *BRCA 1* and *BRCA 2* gene were identified by using MEME (https://meme-suite.org/meme), a free web-based motif elicitation tool. All settings were kept at default, except for number of motifs which was changed to 10 and expected motif sites was set to Any number of repetitions (Anr).

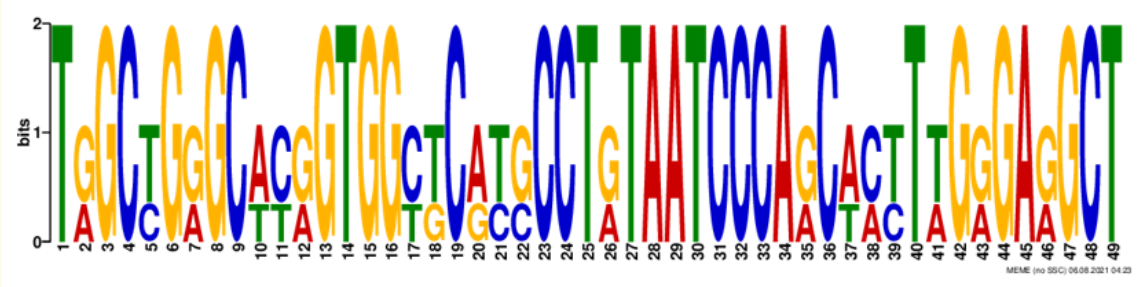

Figure 7: Conserved motif in *BRCA 1* Gene- Height of amino acid represents conservation of the residue.

Figure 8: Conserved motifs and their consensus sequenced (*BRCA 1* Gene).

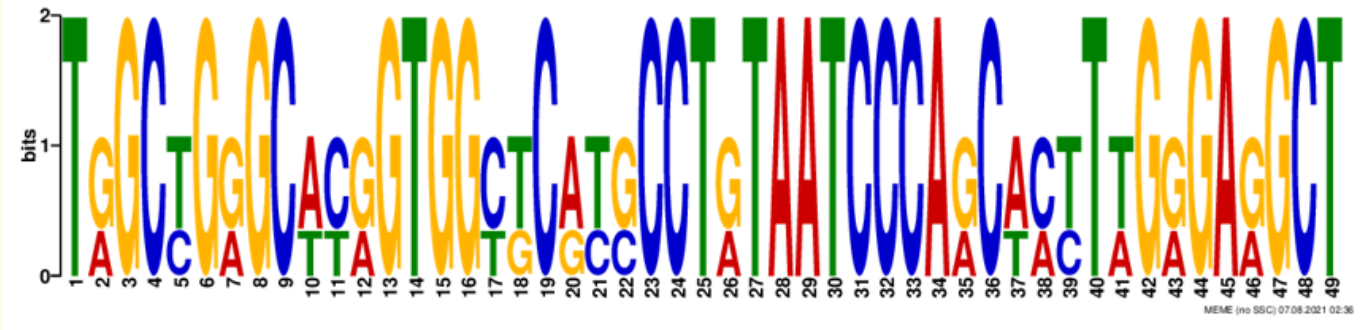

Figure 9: Conserved motif in *BRCA 2* Gene- Height of amino acid represents conservation of the residue

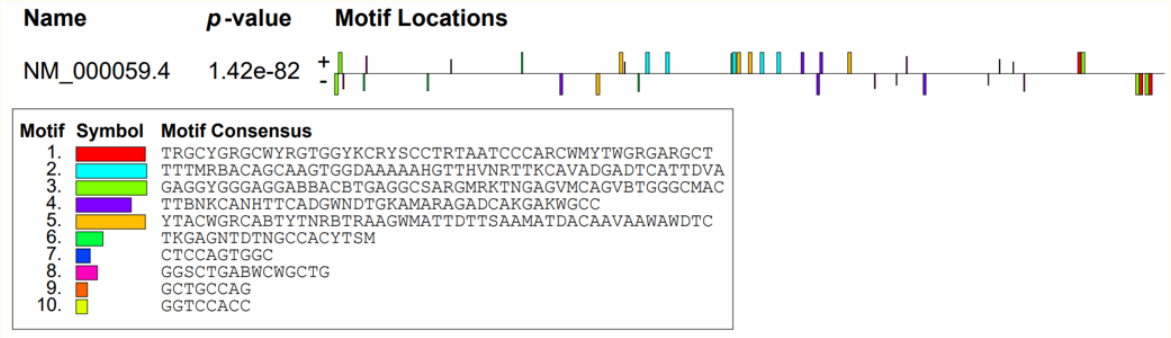

Figure 10: Conserved motifs and their consensus sequenced (*BRCA 2* Gene).

Ontology analysis of *BRCA 1* and *BRCA 2* gene

BRCA 1 gene encodes a 190KDalton nuclear phosphoprotein which has been reported to play an essential role in regulation of genomic stability and is also a vital tumor suppressor. As mentioned earlier, it contains 22 exons (110kb). The BRCA protein forms a multi-subunit complex after associating with various tumor suppressors, signal transducers and DNA damage sensing proteins to form *BRCA1*-Associated genome Surveillance Complex (BASC). This complex can associate with RNA Pol II; therefore, play vital role in transcription, recombination, and DNA repair. Forty percent of inherited breast cancer and 80% of breast and ovarian cancer is said to be caused due to deleterious mutations in *BRCA 1* gene. In table 1 major 40 genes that cause high level of mutational expression and also mention the proteins that form from these genes.

Gene ID	Symbol	Name	Type	# Mutations	Annotations
ENSG00000141510	TP53	tumor protein p53	protein_coding	240	Cancer Gene Census
ENSG00000121879	PIK3CA	phosphatidylinositol-4,5-bisphosphate 3-kinase, catalytic subunit alpha	protein_coding	82	Cancer Gene Census
ENSG00000039068	CDH1	cadherin 1, type 1, E-cadherin (epithelial)	protein_coding	133	Cancer Gene Census
ENSG00000107485	GATA3	GATA binding protein 3	protein_coding	107	Cancer Gene Census
ENSG00000055609	KMT2C	lysine (K)-specific methyltransferase 2C	protein_coding	128	Cancer Gene Census
ENSG00000095015	MAP3K1	mitogen-activated protein kinase kinase kinase 1, E3 ubiquitin protein ligase	protein_coding	134	Cancer Gene Census
ENSG00000171862	PTEN	phosphatase and tensin homolog	protein_coding	66	Cancer Gene Census
ENSG00000141027	NCOR1	nuclear receptor corepressor 1	protein_coding	60	Cancer Gene Census
ENSG00000196712	NF1	neurofibromin 1	protein_coding	64	Cancer Gene Census
ENSG00000159216	RUNX1	runt-related transcription factor 1	protein_coding	46	Cancer Gene Census
ENSG00000117713	ARID1A	AT rich interactive domain 1A (SWI-like)	protein_coding	47	Cancer Gene Census

ENSG00000145675	PIK3R1	phosphoinositide-3-kinase, regulatory subunit 1 (alpha)	protein_coding	47	Cancer Gene Census
ENSG00000065559	MAP2K4	mitogen-activated protein kinase kinase 4	protein_coding	41	Cancer Gene Census
ENSG00000173821	RNF213	ring finger protein 213	protein_coding	57	Cancer Gene Census
ENSG00000167548	KMT2D	lysine (K)-specific methyltransferase 2D	protein_coding	50	Cancer Gene Census
ENSG00000065526	SPEN	spen family transcriptional repressor	protein_coding	60	Cancer Gene Census
ENSG00000083857	FAT1	FAT atypical cadherin 1	protein_coding	43	Cancer Gene Census
ENSG00000127914	AKAP9	A kinase (PRKA) anchor protein 9	protein_coding	53	Cancer Gene Census
ENSG00000100345	MYH9	myosin, heavy chain 9, non-muscle	protein_coding	41	Cancer Gene Census
ENSG00000184634	MED12	mediator complex subunit 12	protein_coding	39	Cancer Gene Census
ENSG00000135111	TBX3	T-box 3	protein_coding	42	Cancer Gene Census
ENSG00000118058	KMT2A	lysine (K)-specific methyltransferase 2A	protein_coding	41	Cancer Gene Census
ENSG00000130396	MLLT4	myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila); translocated to, 4	protein_coding	33	Cancer Gene Census
ENSG00000127329	PTPRB	protein tyrosine phosphatase, receptor type, B	protein coding	35	Cancer Gene Census
ENSG00000129514	FOXA1	forkhead box A1	protein coding	36	Cancer Gene Census
ENSG00000085224	ATRX	alpha thalassemia/mental retardation syndrome X-linked	protein coding	54	Cancer Gene Census
ENSG00000149311	ATM	ATM serine/threonine kinase	protein coding	36	Cancer Gene Census
ENSG00000047410	TPR	translocated promoter region, nuclear basket protein	protein coding	36	Cancer Gene Census
ENSG00000111642	CHD4	chromodomain helicase DNA binding protein 4	protein coding	36	Cancer Gene Census
ENSG00000104517	UBR5	ubiquitin protein ligase E3 component n-recognin 5	protein coding	35	Cancer Gene Census
ENSG00000139618	BRCA2	breast cancer 2, early onset	protein coding	33	Cancer Gene Census
ENSG00000181555	SETD2	SET domain containing 2	protein coding	35	Cancer Gene Census
ENSG00000012048	BRCA1	breast cancer 1, early onset	protein coding	29	Cancer Gene Census

ENSG00000198793	MTOR	mechanistic target of rapamycin (serine/threonine kinase)	protein coding	30	Cancer Gene Census
ENSG00000139687	RB1	retinoblastoma 1	protein coding	32	Cancer Gene Census
ENSG00000079102	RUNX1T1	runt-related transcription factor 1; translocated to, 1 (cyclin D-related)	protein coding	30	Cancer Gene Census
ENSG00000196159	FAT4	FAT atypical cadherin 4	protein coding	32	Cancer Gene Census
ENSG00000138336	TET1	tet methylcytosine dioxygenase 1	protein coding	34	Cancer Gene Census
ENSG00000175054	ATR	ATR serine/threonine kinase	protein coding	32	Cancer Gene Census
ENSG00000178104	PDE4DIP	phosphodiesterase 4D interacting protein	protein coding	31	Cancer Gene Census

Table 1: Highly mutated genes and their protein production.

Inherited mutations in *BRCA1* and this gene, *BRCA2*, confer increased lifetime risk of developing breast or ovarian cancer. Both *BRCA1* and *BRCA2* are involved in maintenance of genome stability, specifically the homologous recombination pathway for double-strand DNA repair. The largest exon in both genes is exon 11, which harbors the most important and frequent mutations in breast cancer patients. The *BRCA2* gene was found on chromosome 13q12.3 in human. The *BRCA2* protein contains several copies of a 70 aa motif called the BRC motif, and these motifs mediate binding to the RAD51 recombinase which functions in DNA repair. *BRCA2* is considered a tumor suppressor gene, as tumors with *BRCA2* mutations generally exhibit loss of heterozygosity (LOH) of the wild-type allele.

Novel methods for mutational disease prediction

The major cause of deaths in the world is cancer. In 2004, 7.4 million deaths were reported and it is thought that in 2030, this rate will reach near the 12 million deaths [24]. The second major reason behind cancer deaths globally is the breast cancer. The abnormal division and growth of cells in the breast tissue leads towards the malignant tumor [22]. There is need to discover efficient methods of computer aided diagnosis for the treatment and prevention of cancers. Currently, biologists are using the computers just like highly skilled professionals and solving the multiple problems that are very specific to the biologist by using the tools of bioinformatics. The purpose of using the bioinformatics tools is to explain the functioning of living organisms and to examine the biological statistics [23].

Mainly, two methods are being used for the prediction of mutational diseases; first is to know that if the mutations are causing the disease in the patients or not, and the second is to determine the relation of certain diseases with these mutations.

Detection of differentially expressed mutant genes of *BRCA1/2*

There is need of genetic testing of cancer patients to check that either their families have cancer history or not. Somatic and germline variants are considered to be associated with the risk of cancer to some extent (Foulkes, 2013; Shimada et al., 2019). For the treatment and diagnosis of breast cancer, there is need to find more and more genes involved in the disease. The Cancer Genome Atlas (TCGA) database is a very helpful approach for researchers in the identification of relationship between disease and genetic molecules. From the TCGA database (https://portal.gdc.cancer.gov/), the clinical profiles of over 1092 human breast cancer patients were downloaded. The cBioPortal website (http://www.cbioportal.org/index.do) was used for obtaining information related to the mutations of *BRCA1/2*. From the whole breast cancer RNA-sequence data, the wild-type (WT) and mutant-type (MT) were selected indiscriminately. As a control group, whole correspondent para-carcinoma tissue samples were selected from breast cancer RNA-sequence statistics.

The Gene Set Enrichment Analysis (GSEA) was used to examine the modifications of two groups either with mutations or without

mutations in BRCA genes. The transformations figure was set at the same number with the default parameters. The gene ID transformation and default significance amputation were used to manage the expression profile data of the examined collections. The protein-protein interaction (PPI) construction helped in empathetic the association of the genetic expressions between *BRCA1/2* wild-type breast cancer patients and mutant breast cancer patients. For the screening of differentially expressed genes in all comparisons, Kaplan-Meier plotter (http://kmplot.com/analysis/), an online software was used.

Results and Discussion

According to many researches, *BRCA1* and *BRCA2* genes have been generally recognized as the major penetrant genes that cause a genetic susceptibility to breast cancer syndrome. Structure analysis performed in this study which revealed that *BRCA1* has 24 exons and *BRCA2* has 27 exons are located on chromosome 17 and 13 respectively. Pfam and SMART shows that, *BRCA1* has 13 conserved regions, 3 shows uniqueness to humans and 21 likely disease associated missense changes, 11 of which occur in conserved regions, and *BRCA2* has longest internal conserved domain encoded by exons 14-24. BRCA genes have multiple domains, *BRCA1* has 4 and *BRCA2* has 12 domains. According to the results shown by PTM code 2, *BRCA1* has 50PTMs and *BRCA2* has 95 PTMs annotations. The result from STRING 9.1 was showing the association of BRCA genes. BRCA genes are closely link with several different proteins, ESR1, BRIPI1, H2AFX, BARD1,RBBP8,TP53,UMC1, RAD51, and ATM. By using BioEdit as a new approach, mutation can be detecting in TP53 of BRCA genes.

By using the TCGA database, we were able to collect comprehensive breast cancer clinical profiles and the RNA-Seq dataset that accompanied them. According to the cBioPortal website, 6–10% of BC patients had *BRCA1/2* mutations, while the rest had *BRCA1/2* wild type; among them, the proportion of *BRCA1* and *BRCA2* mutations was 3 percent (38/1098, TCGA: Provisional) and 4 percent (48/1098, TCGA: Preliminary) (or, 4 percent, 45/1088 and 5 percent, 54/1088, in TCGA: PanCancer Atlas). Missense mutations and frame-shifts were the most common mutation types in *BRCA1* and *BRCA2* genes as shown in table 2.

DNA Change	Type	Consequences	# Affected Cases in Cohort	Impact
chr3:g.179234297A>G	Substitution	Missense PIK3CA H1047R	121/981,12.33%	VEP: MODERATE, SIFT: tolerated - score 0.11, PolyPhen: possibly_damaging - score 0.529
chr3:g.179218303G>A	Substitution	Missense PIK3CA E545K	63/981,6.42%	VEP: MODERATE, SIFT: deleterious - score 0.02, PolyPhen: probably_damaging - score 0.959
chr3:g.179218294G>A	Substitution	Missense PIK3CA E542K	43/981,4.38%	VEP: MODERATE, SIFT: deleterious - score 0.04, PolyPhen: probably_damaging - score 0.96
chr14:g.104780214C>T	Substitution	Missense AKT1 E17K	25/981,2.55%	VEP: MODERATE, SIFT: deleterious - score 0, PolyPhen: probably_damaging - score 0.999
chr10:g.8069470delCA	Deletion	Splice Acceptor GATA3 X309_splice	21/981,2.14%	VEP: HIGH
chr17:g.7675088C>T	Substitution	Missense TP53 R175H	20/981,2.04%	VEP: MODERATE, SIFT: tolerated - score 0.11, PolyPhen: benign - score 0.308
chr3:g.179203765T>A	Substitution	Missense PIK3CA N345K	17/981,1.73%	VEP: MODERATE, SIFT: deleterious - score 0.01, PolyPhen: probably_damaging - score 1

chr3:g.179234297A>T	Substitution	Missense PIK3CA H1047L	13/981,1.33%	VEP: MODERATE, SIFT: tolerated - score 0.44, PolyPhen: benign - score 0.085
chr17:g.7673802C>T	Substitution	Missense TP53 R273H	12/981,1.22%	VEP: MODERATE, SIFT: tolerated - score 0.13, PolyPhen: possibly_damaging - score 0.631
chr2:g.197402110T>C	Substitution	Missense SF3B1 K700E	10/981,1.02%	VEP: MODERATE, SIFT: deleterious - score 0, PolyPhen: probably_damaging - score 0.96
chr3:g.179221146G>A	Substitution	Missense PIK3CA E726K	9/981,0.92%	VEP: MODERATE, SIFT: tolerated - score 0.36, PolyPhen: benign - score 0.4
chr6:g.167976333_167976334insA	Insertion	Downstream Gene MLLT4	9/981,0.92%	VEP: MODIFIER
chr16:g.68738315C>T	Substitution	Stop Gained CDH1 Q23*	9/981,0.92%	VEP: HIGH
chr10:g.8069550_8069551insG	Insertion	Frame-shift GATA3 D336Gfs*17	8/981,0.82%	VEP: HIGH
chr17:g.7674945G>A	Substitution	Stop Gained TP53 R196*	8/981,0.82%	VEP: HIGH
chr17:g.7670685G>A	Substitution	Stop Gained TP53 R342*	7/981,0.71%	VEP: HIGH
chr3:g.179210291G>A	Substitution	Missense PIK3CA E453K	7/981,0.71%	VEP: MODERATE, SIFT: tolerated - score 0.06, PolyPhen: possibly_damaging - score 0.801
chr17:g.7674872T>C	Substitution	Missense TP53 Y220C	7/981,0.71%	VEP: MODERATE, SIFT: deleterious - score 0, PolyPhen: probably_damaging - score 1
chr17:g.7674894G>A	Substitution	Stop Gained TP53 R213*	6/981,0.61%	VEP: HIGH
chr17:g.7674953T>C	Substitution	Missense TP53 H193R	6/981,0.61%	VEP: MODERATE, SIFT: deleterious - score 0, PolyPhen: probably_damaging - score 1
chr3:g.179218307A>G	Substitution	Missense PIK3CA Q546R	6/981,0.61%	VEP: MODERATE, SIFT: deleterious - score 0.02, PolyPhen: probably_damaging - score 0.984
chr17:g.7675076T>C	Substitution	Missense TP53 H179R	6/981,0.61%	VEP: MODERATE, SIFT: deleterious - score 0, PolyPhen: probably_damaging - score 1
chr10:g.8073911_8073912insG	Insertion	Frame-shift GATA3 P409Afs*99	6/981,0.61%	VEP: HIGH

Table 2: Major DNA changes involves in mutations for causing breast cancer.

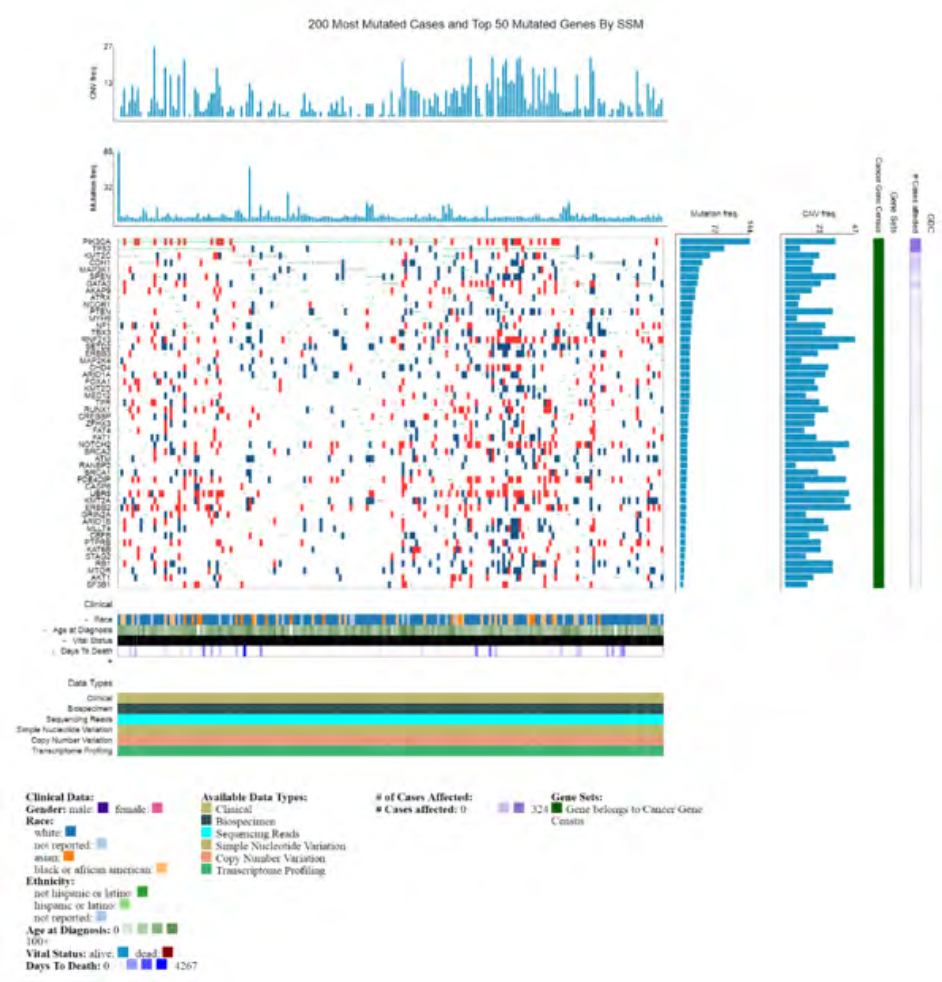

Figure 11: OncoGrid, 200 Most Mutated Cases and Top 50 Mutated Genes By SSM.

Compared to the 110 wild type group, the 50 mutant groups (as in figure 11, oncogrid shows top 50 mutated genes in 200 mutated cases) had considerably more of the seven, eighty-three, or one conventional gene sets from the Hallmark collection and the c2 KEGG-sub collection and c5 collection, respectively. There was a clear enrichment of gene sets related to the mitotic spindle, cell cycle, G2M checkpoint, and others. As a result of DEG screening, RNA-Seq datasets from 50 *BRCA1/2* mutation-carrier patients and 110 wild-type BC patients were utilized in total. Two hundred and ninety-four of the 297 DEGs found between BRCA-mutant and wild-type BC were up-regulated, whereas 202 were down-regulated. On the other hand, a comparison of the WT (breast cancer) and control group tissues revealed 5255 genes that were expressed differently in BC patients. Certain molecular processes of BRCA mutations' carcinogenesis may be influenced by these genes, according to research. Table 3 lists the top five overlapping DEGs that were up-and-down-regulated based on fold changes.

The genes CCNE1, MAOA, NPBWR1, EXO1, SLC4A4, A2ML1, and TTK (dual-specificity protein kinase) were shown to be

Top 5 up-regulated genes in *BRCA1/2*-mutant breast cancer			
	MT vs WT	MT vs C	WT vs control
CT45A10	4.46	7.66	3.22
TBX10	3.94	6.63	2.61
BARHL2	3.24	5.40	2.15
NLRP7	3.26	4.97	1.71
CLLU1OS	2.68	4.15	1.47
Top 5 down-regulated genes in *BRCA1/2*-mutant breast cancer			
CA4	−3.11	−8.48	−5.35
MYOM2	−3.24	−5.64	−2.42
SLC4A4	−2.07	−3.48	−1.37
LGALS12	−2.41	−6.47	−4.04
HPSE2	−1.90	−5.12	−3.23
CAPN11	−1.91	−3.23	−1.36

Table 3: Top five overlapping DEGs that were up-and-down-regulated.

MT* mutant type, WT* wild type and C* control.

substantially upregulated in *BRCA1/2*-mutant breast cancer compared to WT breast cancer and normal tissue. Accordingly, we believed that the expression of CCNE1, TTK, and EXO1 in BC tissues was significantly higher than in para-carcinoma tissues, and that this could be a promising way to screen BC and then further differentiate the high risks of *BRCA1/2* mutations from wild-type BC, while neglecting the potential effect of the genetic background associated to the ethnic background.

We discovered some key compounds impacted by BRCA gene mutations using bioinformatic studies. Genes' potential as reliable prognostic or diagnostic markers, as well as prospective therapeutic targets, were suggested by survival and diagnosis analyses, as well as the validation of genes in ethnic subgroups. A2ML1, EXO1, CCNE1, and TTK were identified as novel DEGs that may play key roles in the oncogenesis and development of *BRCA1/2*-mutant BC. The *BRCA1/2*-mutant individuals in this study are taken from TCGA cancer patients who do not differentiate among somatic and germline abnormalities (not necessarily as hereditary breast cancer patients). The identification of hub genes was also confirmed in additional analyses of mutant and standard tissues, as well as wild-type and standard tissues, to pick more credible important candidate genes. To sum it up, our work will shed light on the role of BRCA in carcinogenesis, uncover new diagnostic markers for breast cancer and *BRCA1/2*-mutant breast cancer, as well as give new targets and techniques for individualized therapy.

New approach for prediction pre-cancer via detecting mutated in tumor Protein P53

Tumor P53 protein is formed by using TP53 gene. The cell division is regulated by P53 tumor protein and acts as a tumor suppressor. This protein is present in the nucleus of cells and connected to the DNA directly. On the short arm of chromosome number 17, the TP53 gene is present at position 13 [20].

The chances of cancer increases as the age increases. According to the recent studies, it is revealed that there are mutations in the gene P53 observed in 50% cancer patients therefore P53 is an important gene in cancer study [19]. The normal TP53 gene and its protein is used along with bioinformatics tools as a new method for prediction of pre-cancer mutations in tumor protein P53 [20]. The TP53 is required to categorize and analyze the malignant mutations of TP53 because without using the TP53 database, it is difficult to detect various types of cancers. The BioEdit ver. 7.2.0, FASTA, BLAST, CLUSTALW, and TP53 databases were used for the prediction of pre-cancer.

Conclusion

The *BRCA1* and *BRCA2* genes defend against some cancers. However, some abnormalities in the *BRCA1* and *BRCA2* genes prevent them from functioning correctly, making you more likely to somehow get breast, ovarian, and other cancers if you inherit one of these mutations. Hence it is concluded that by using bioinformatics tools, it is possible to analyze, locate, and detect conserved regions/domains of several mutational diseases.

Bibliography

1. Miki Yoshio., *et al*. "A strong candidate for the breast and ovarian cancer susceptibility gene *BRCA1*". *Science* 266.5182 (1994): 66-71.

2. Li Yue., *et al*. "Differentially expressed genes and key molecules of *BRCA1/2*-mutant breast cancer: evidence from bioinformatics analyses". *Peer Journal* 8 (2020): e8403.

3. Wooster Richard., *et al*. "Identification of the breast cancer susceptibility gene *BRCA2*". *Nature* 378.6559 (1995): 789-792.

4. Antoniou Antonis C., *et al*. "Breast and ovarian cancer risks to carriers of the *BRCA1* 5382insC and 185delAG and *BRCA2* 6174delT mutations: a combined analysis of 22 population based studies". *Journal of Medical Genetics* 42.7 (2005): 602-603.

5. Chen Sining and Giovanni Parmigiani. "Meta-analysis of *BRCA1* and *BRCA2* penetrance". *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology* 25.11 (2007): 1329.

6. Mavaddat Nasim., *et al*. "Cancer risks for *BRCA1* and *BRCA2* mutation carriers: results from prospective analysis of EMBRACE". *JNCI: Journal of the National Cancer Institute* 105.11 (2013): 812-822.

7. Wooster Richard., *et al*. "Identification of the Breast-Cancer Susceptibility Gene *BRCA2* (Vol 378, Pg 789, 1995)". *Nature* 379.6567 (1996): 749-749.

8. Chapman Mark S and Inder M Verma. "Transcriptional activation by *BRCA1*". *Nature* 382.6593 (1996): 678-679.

9. Kolinjivadi Arun Mouli., *et al*. "Moonlighting at replication forks–a new life for homologous recombination proteins *BRCA 1*, *BRCA 2* and RAD 51". *FEBS Letters* 591.8 (2017): 1083-1100.

10. Chen Junjie., *et al*. "Stable interaction between the products of the *BRCA1* and *BRCA2* tumor suppressor genes in mitotic and meiotic cells". *Molecular Cell* 2.3 (1998): 317-328.

11. Rebbeck Timothy R., *et al*. "Mutational spectrum in a worldwide study of 29,700 families with *BRCA1* or *BRCA2* mutations". *Human Mutation* 39.5 (2018): 593-620.

12. Smith SA., *et al*. "Allele losses in the region 17q12–21 in familial breast and ovarian cancer involve the wild–type chromosome". *Nature Genetics* 2.2 (1992): 128-131.

13. Thull Darcy L and Victor G Vogel. "Recognition and management of hereditary breast cancer syndromes". *The Oncologist* 9.1 (2004): 13-24.

14. Alvarez Carolina., *et al*. "Different Array CGH profiles within hereditary breast cancer tumors associated to *BRCA1* expression and overall survival". *Bmc Cancer* 16.1 (2016): 1-14.

15. Wang Zhu., *et al*. "Expression and mutations of BRCA in breast cancer and ovarian cancer: Evidence from bioinformatics analyses". *International Journal of Molecular Medicine* 42.6 (2018): 3542-3550.

16. Lee Megan S., *et al*. "Comprehensive analysis of missense variations in the BRCT domain of *BRCA1* by structural and functional assays". *Cancer Research* 70.12 (2010): 4880-4890.

17. Monteiro Alvaro N., *et al*. "Variants of uncertain clinical significance in hereditary breast and ovarian cancer genes: best practices in functional analysis for clinical annotation". *Journal of Medical Genetics* 57.8 (2020): 509-518.

18. Guidugli Lucia., *et al*. "A classification model for *BRCA2* DNA binding domain missense variants based on homology-directed repair activity". *Cancer Research* 73.1 (2013): 265-275.

19. Carlsson Jonas Thierry Soussi and Bengt Persson. "Investigation and prediction of the severity of p53 mutants using parameters from structural calculations". *The FEBS journal* 276.15 (2009): 4142-4155.

20. Ismaeel Ayad Ghany. "New Approach for Prediction Pre-cancer via Detecting Mutated in Tumor Protein P53". *arXiv preprint arXiv:1310.2182* (2013).

21. Ismaeel Ayad Ghany and Anar Auda Ablahad. "Novel method for mutational disease prediction using bioinformatics techniques and backpropagation algorithm". *arXiv preprint arXiv:1303.0539* (2013).

22. Kharya Shweta. "Using data mining techniques for diagnosis and prognosis of cancer disease". *arXiv preprint arXiv:1205.1923* (2012).

23. Claverie Jean-Michel and Cedric Notredame. Bioinformatics for dummies. John Wiley and Sons (2006).

24. Deepa R., *et al*. "Bio-Microbe Analyst: New-Fangled Unambiguous Syndrome Prediction to Classify Cancer with Recent Computational Gist" (2013).