Targeted Gene Panel Sequencing Unveiled New Pathogenic Mutations in Patients With Breast Cancer

Bioinformatics and Biology Insights
Volume 17: 1–10

DOI: 10.1177/11779322231182054

Souad Kartti[1,2*], El Mehdi Bouricha[1,2*], Oumaima Zarrik[1], Youssef Aghlallou[3], Chaimaa Mounjid[4], Rachid ELJaoudi[1,2,5], Lahcen Belyamani[2,5,6], Azeddine Ibrahimi[1,2,6] and Basma EL khannoussi[4]

[1]Biotechnology Lab (MedBiotech), Bioinova Research Center, Rabat Medical & Pharmacy School, Mohammed V University in Rabat, Rabat, Morocco. [2]Mohammed VI Center for Research and Innovation, Rabat, Morocco. [3]Institute for Research in Cancer, Fes, Morocco. [4]Pathology Department, Oncology National Institute, Rabat Medical and Pharmacy School, Mohammed V University in Rabat, Rabat, Morocco. [5]Emergency Department, Military Hospital Mohammed V, Rabat, Morocco. [6]Mohammed VI University of Health Sciences, Casablanca, Morocco.

ABSTRACT: The increasing commercialization of new gene panels based on next-generation sequencing for clinical research has significantly improved our understanding of breast cancer genetics and has led to the discovery of new mutation variants. The study included 16 unselected Moroccan breast cancer patients tested with multi-gene panel (HEVA screen panel) using Illumina Miseq, followed by Sanger sequencing to validate the most relevant mutation. Mutational analysis revealed the presence of 13 mutations (11 single-nucleotide polymorphisms [SNPs] and 2 indels), and 6 of 11 identified SNPs were predicted as pathogenic. One of the 6 pathogenic mutations was c.7874G>C, a heterozygous SNP in HD-OB domain of BRCA2 gene, which led to the arginine to threonine change at codon 2625 of the protein. This work describes the first case of a patient with breast cancer harboring this pathogenic variant and analyzes its functional impact using molecular docking and molecular dynamics simulation. Further experimental investigations are needed to validate its pathogenicity and to verify its association with breast cancer.

KEYWORDS: Multi-gene panel, breast cancer, *BRCA2*, next-generation sequencing, c.7874G>C, molecular dynamics

RECEIVED: March 27, 2023. **ACCEPTED:** May 27, 2023.

TYPE: Original Research Article

FUNDING: The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The Institute of Cancer Research in Morocco provided funding for this study through grant No. 591/AAP 2017.

DECLARATION OF CONFLICTING INTERESTS: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

CORRESPONDING AUTHOR: Souad Kartti, Mohammed VI Center for Research and Innovation, 10112 Rabat, Morocco. Email: souad.kartti@um5s.net.ma

Introduction

Breast cancer (BC) is the major cause of cancer deaths in women and the most diagnosed malignancy in the world.[1] The incidence of BC in North Africa was higher in the continent with 29.3 per 100 000 (95% confidence interval [CI] = 20.0-38.7).[2] The GLOBOCAN 2020 database of the International Agency for Research on Cancer (IARC) estimated that in Morocco BC represents 19.8 % of all cancer cases. Newly diagnosed BC cases were 38.9 % of female cancers in 2020; it was responsible for 10.5% of deaths in the country and it represents the second cause of death after lung cancer.

Breast cancer is believed to be a complicated disease that arises from a combination of environmental, microbiome, and genetics factors that increase its pathogenicity.[3] Indeed, at the genetics level, BC is associated with a number of genes with high- and moderate to low-penetrance susceptibility, implicated in DNA damage response. Breast cancer type 1 susceptibility protein (BRCA1) located on chromosome 17 and breast cancer type 2 susceptibility protein (*BRCA2*) located on chromosome 13 are the most important genes associated with BC, with a 60% and 55%, respectively, cumulative risks.[4] Mutations in other genes such as *PTEN*, *CDH1*, *PALB2*, *STK11*, and *ATM* were proven to confer an increased risk of BC.[5] Especially, *PALB2* and *PTEN* are considered high-risk genes with a lifetime risk above 40%. On the other side, there are BC genes with low/moderate risk like NBN.[6] Furthermore, studies have identified a number of additional DNA repair genes that interact with *BRCA1*, *BRCA2*, and/or the BRCA pathways and confer about a 2-fold increase in BC risk, including *CHEK2*,[7] *BRIP1* (*BACH1*),[8] *ATM*,[9] and *PALB2*.[10] Moreover, a mutation in a tumor suppressor gene, such as *TP53*, *BRCA1*, *BRCA2*, *CHEK2*, *PTEN*, or *ATM*, can be sporadic or germinal, which results in the promotion of breast carcinogenesis mechanism. Inherited mutations in certain genes such as TP53, BRCA1, and BRCA2 have been linked to an increased risk of BC. Similarly, sporadic mutations in TP53 have been detected in BC cells. These genes are crucial in maintaining genomic stability, and any mutations that affect their function can hamper the processing of DNA damage. This is likely the primary mechanism by which mutations in these 5 genes increase the risk of BC.[11] Germline pathogenic variants in high-risk genes allowed to prevent the onset of cancer and to undertake risk-reduction strategies with greater efficiency, although germline pathogenic variants detected in low- or moderate-risk BC genes have also a significant clinical impact.[12,13] In addition, it has been shown that many variants may contribute to BC

* These authors contributed equally.

development, such as *MSH6*, *RAD50*, *RAD51*, *APC*, *BARD1*, *MLH1*, *MSH2*, *MUTYH*, *NBN*, and *PMS2*.[14-18] Furthermore, high-risk BC genes are those with a BC odds ratio >5.0 (*BRCA1*, *BRCA2*, *PALB2*, *STK11*, *TP53*, *PTEN*, and *CDH1*), according to a previous study.[6]

With the advancement of next-generation sequencing (NGS) techniques, multi-gene panel testing is providing a cost-effective way of diagnosis, which allows the identification of genetic variants and novel mutations that have an impact on the cancer process.

In this respect, the aim of our study was to assess further potential mutations that may be related to BC by sequencing and analyzing the main 22 genes related to BC (*ATM*, *APC*, *CHEK2*, *BARD1*, *BRCA1*, *BRCA2*, *MSH2*, *BRIP1*, *CDH1*, *PALB2*, *EPCAM*, *MLH1*, *MSH6*, *MUTYH*, *NBN*, *PMS2*, *PTEN*, *RAD50*, *RAD51C*, *RAD51D*, *STK11* and *TP53*) grouped in the HEVA screen kit. This analysis was conducted on 16 Moroccan patients with BC.

Materials and Methods

Patients

In this study, fresh breast tumor tissue was collected from 16 patients who had been diagnosed with primary BC and had undergone surgery for mastectomy at the National Institute of Oncology, Ibn Sina University Hospital of Rabat, Morocco, between June and December 2020. This study was approved by the Ethical Committee of the Faculty of Medicine and Pharmacy in Rabat, Morocco, with reference number 103/17.

DNA extraction

Genomic DNA was extracted from fresh breast tissue using QIAamp DNA Mini Kit (Qiagen GmbH, Germany). For validation of the observed mutation, DNA from blood was also isolated using the QIAamp DNA Mini Kit Blood kit (Qiagen) according to the specification of the manufacturer. DNA was quantified by Qubit fluorometer 2.0 (Thermo Fisher Scientific, Waltham, MA, USA) and its quality was assessed using 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) and stored at −20°C until analyzed.

Amplicon sequencing gene panel approach

We used a commercial kit HEVA screen (4 bases SA) to analyze 22 genes associated with BC (*ATM*, *APC*, *BARD1*, *BRCA1*, *BRCA2*, *BRIP1*, *CDH1*, *CHEK2*, *EPCAM*, *MLH1*, *MSH2*, *MSH6*, *MUTYH*, *NBN*, *PALB2*, *PMS2*, *PTEN*, *RAD50*, *RAD51C*, *RAD51D*, *STK11*, *TP53*). A molecular protocol based on NGS has been used. We used 20 ng of DNA that was amplified in 3 independent multiplex polymerase chain reactions (PCRs). Then products were pooled for each sample, 1 µL of Reagent A was added, and the samples were incubated at 50°C for 10 minutes, 55°C for 10 minutes, and 60°C for 20 minutes. After that, a ligation mix was added, and samples continued in incubation at 22°C for 30 minutes and 72°C for 10 minutes. Library purification was performed using Agencourt Ampure beads (Beckman Coulter) and then eluted into a volume of 40 µL. Purified libraries underwent the insertion of adapters and indexes in the Index PCR profile and then was again purified with Agencourt Ampure beads (Beckman Coulter). The quantity and quality of prepared libraries were assessed by Qubit 2.0 fluorometer (Thermo Fisher Scientific). Each library was diluted to a concentration of 4 nM according to the manufacturer's instructions and samples were pooled to 1 final library at equimolar ratios. A further denaturation was necessary using 0.2 N NaOH, diluted with HT1 buffer to 12 pM concentration, and 20% of PhiX Control v3 was added (Illumina). Library products were analyzed by massive parallel sequencing using MiSeq Reagent Kit v3 (600-cycle) (Illumina) on the Illumina platform MiSeq in 2 × 251 cycle profile.

Variant calling and genetic annotation

The study processed the raw paired-end FASTQ files using Trimmomatic v0.33 to filter out poor-quality reads and aligned them to the reference genome (GRCh37/hg19) using Burrows-Wheeler Aligner (BWA) v0.7.10. The subsequent steps involved removing PCR duplicates, overrepresented sequences, and low-quality reads. The identification of variants from aligned reads was carried out using the GATK software package v1.6, and coverage analysis was done in parallel with Samtools[19]. The study excluded variants such as single-nucleotide variants (SNVs) and insertions/deletions (InDels) located outside the exon region (except for splicing sites) and synonymous SNVs. Finally, the ANNOVAR software[20] was used to annotate the variants, and the Integrative Genomics Viewer[21] (IGV; Broad Institute, Cambridge, MA, USA) was used to visualize variants against the reference genome.

Gene data analysis

The identified variants were graphically presented using OncoPrint, and Mutation Mapper was implemented in cBioportal (http://cbioportal.org). These tools provided information about genomic alterations and gave a graphical representation of the mutations found in each gene queried, survival analysis, patient-centric queries, and network visualization and analysis.

Mutation validation

Sanger sequencing was performed for tissue and blood DNA samples to confirm the presence of the R2625T mutation and to determine whether the mutation is germline or somatic. The targeted region was amplified using the following primers: *BRCA2*_F: GCCACCATGCTCAGCAATGAA and *BRCA2*_R: GTCACTGACAACTGGCTTGTG; those

primers were designed using NCBI/PrimerBLAST. The PCR reaction volume was 25 µL, composed of MyTaq Mix (Bioline, BIO-25041) 12.5 µL, ddH_2O 7.5 µL, forward and reverse primers 2 µL, and template DNA 3 µL. The PCR reaction program was as follows: pre-denaturation at 95°C for 3 minutes, denaturation at 94°C for 30 seconds, annealing at 60°C for 30 seconds, extension at 72°C for 30 seconds for 30 cycles, final extension at 72°C for 10 minutes, and storage at 4°C. The purified PCR products were sequenced using the ABI3500 Genetic Analyzer (Applied Biosystems) and the BigDye X-terminator v3.1 Cycle Sequencing Kit (catalog number 4376486; Applied Biosystems) according to the supplier's protocol. Integrative Genomics Viewer software[22] and chromatogram visualization were used to analyze and visualize the R2625T mutation.

Functional protein prediction

The functional impact of mutations was evaluated using 4 different bioinformatics algorithms: SIFT (Sorting Intolerant From Tolerant),[23] PolyPhen-2 (PP-2), MutationTaster,[24] and LTRpred.[25] A SNP was considered potentially pathogenic if it was classified as deleterious by all 4 algorithms.

3D structure modeling of helical and OB1 domain of BRCA2

The 3-dimensional (3D) model of helical domain (HD) and OB1 domain (amino-acid sequence: residues 2481-2832) of BRCA2 was generated using AlphaFold software. The Mutant model was built by making point mutations R2625T on the generated AlphaFold model with the Rotamers tool of UCSF Chimera.[26] Then, the mutant model was relaxed by 1000 steps of SD followed by 1000 steps of CG minimizations, keeping all atoms far by more than 5Å from the mutated residue fixed.

Molecular docking between HD-OB1 and DSS1

To gain insight into the impact of mutation on the BRCA2 HD-OB1 and deleted in split hand/split foot protein 1 (DSS1) complex, a docking study was done using HDOCK web server.[27] This server supports protein-protein docking and automatically predicts their interaction through a hybrid algorithm of template-based and template-free docking. For the BRCA2 HD-OB1, the wild-type (WT) and mutated-type (MT) models generated by Alphafold were used. For DSS1, we used the *Homo sapiens* model extracted from 1MIU which was in complex with mouse BRCA2. The resulting complexes were analyzed by the Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) method using the HawkDock server to predict the global binding free energies of the protein-protein systems and to highlight the key residues in the binding interface with the per residue MM/GBSA free energy decomposition function.[28] To validate our docking protocol, 2 known pathogenic mutations located at BRCA2 HD-OB1 were used as reference. These mutations were W2626C and I2627F; the last one is known to affect the interaction between BRCA2 and DSS1 complex by reducing significantly the affinity, while W2626C has been reported to have no effect on the affinity of the complex.[29]

Molecular dynamics simulation

The molecular dynamics (MD) simulation was done for the WT and 3 MT (R2526T, W2626C and I2627F) helical HD-OB1 models in complex with DSS1. Molecular dynamics was carried out for 50 ns with a recording interval of 100 ps using Desmond module.[28] In system builder, the OPLS3e force field was selected and TIP3P was used as a solvent model with a 10Å orthorhombic box and then the system charge was neutralized by adding 0.15 M of sodium (Na+) and chloride ions (Cl^-). The generated system was subjected to energy minimization and equilibrated via NPT ensemble at a constant temperature of 310 K and 1.01325 bar pressure. All other Desmond parameters were kept at their default values. Once MD simulation was done, simulation trajectories were analyzed using the simulation interaction diagram included in Desmond to obtain root mean square deviation (RMSD) and root mean square fluctuation (RMSF) data.

Results

In this study, we evaluated the mutational spectrum of 22 genes among 16 Moroccan BC patients using the HEVA screen panel. Among the 16 patients, 2 were filtered out due to their poor DNA sequencing quality. According to hormonal receptors status (progesterone receptor [PR], estrogen receptor [ER]) and human epidermal growth factor receptor 2 (HER2) of the remaining 14 patients, 2 were ER+/PR+/HER2+, 2 were ER+/PR+/HER2 equivocal, 6 were ER+/PR+/HER2–, 1 was ER+/HER+, and 3 were triple-negative breast cancer (TNBC) (Table 1).

Mutational screening

A total of 11 SNPs and 2 insertions were identified from the targeted sequencing regions of 12 of 14 patients. According to the Oncoprint analysis, the SNP alterations comprised missense and splice site mutations. TP53 emerged as the gene with the highest frequency of mutations, with 4 SNPs, followed by *APC* with 3 SNPs, while *BRCA2*, *CHECK2*, *RAD50*, and *RAD51C* harbored 1 SNP each. For *NBN* and *ATM* genes, only 1 insertion was detected (Figure 1A).

Among the 13 identified variants, only 6 were known and found in different human germinal/somatic variant databases. This includes all the 4 TP53 mutations (c.528C>A, c.785G>T, c.853G>A and c.991C>T), identified, respectively, with cosmic and/or rs numbers (COSM43734, COSM11198, [rs112431538, COSM10722], and COSM11354), the c.8044G>A APC

Table 1. Characteristics of patients with breast cancer stratified by age and hormone receptor status.

PATIENT	AGE	HORMONAL RECEPTOR STATUS		HER2	KI67 (%)
		ER	PR		
1	79	+	+	–	18
2	72	+	+	Equivocal	–10
3	55	+	+	–	–1
4	53	–	–	–	NA
5	47	+	+	+	NA
6	59	–	–	–	NA
7	65	+	+	–	NA
8	37	+	+	+	NA
9	33	+	–	–	NA
10	65	+	+	–	30
11	52	–	–	–	NA
12	42	+	–	+	70
13	49	+	+	–	NA
14	41	+	+	Equivocal	NA

Abbreviations: ER, estrogen receptor; HER2, human epidermal growth factor receptor 2; PR, progesterone receptor.

mutation (COSM4559410), and the c.7874G>C *BRCA2* SNP (COSM7403374) (Table 2).

To evaluate the impact of the 11 SNPs in terms of pathogenicity, we used, in a preliminary step, 4 functional prediction webservers (Sift, Polyphen2, MutationTaster, and LTRpred). The obtained results show that 6 mutations have been predicted as pathogenic by the 4 servers, including 2 mutations of *APC* (c.2965G>A and c.8044G>A), 2 of *TP53* (c785G>T and c.853G>A), and 1 mutation of *BRCA2* (c.7874G>C), while the c.7748A>G of *APC* and c.2776A>G of *RAD50* were predicted as pathogenic by only 3 servers. The A580C was predicted as tolerant (Table 2). The stopgain SNPs at *TP53* (c.528C>A, c.991C>T) and *RAD51C* (c.196A>T) as well as the stopgain insertion at *ATM* (7543-7544 insTCACCGACTGCCCATAGAG) are considered pathogenic as they directly affect the length of the protein. The *NBN* frameshift insertion (1675dupG) was also considered deleterious as it changes the reading frame. Subsequently, Mutation Mapper was used to map the identified mutations on linear proteins and their domains to determine whether these mutations affect the functional domain of proteins (Figure 1B). The results show that 6 mutations are located in functional domains, namely those of *TP53* (G262V, E285K, C176*, Q331*) which are situated in the DNA-binding domain (DBD) and the P53 tetramerization domain, while the V2682M mutation of *APC* gene is located at the EB-1-binding domain and finally the R2625T mutation of BRCA2 is located at the helical domain (HD-OB1). This last mutation has been cited in a previous

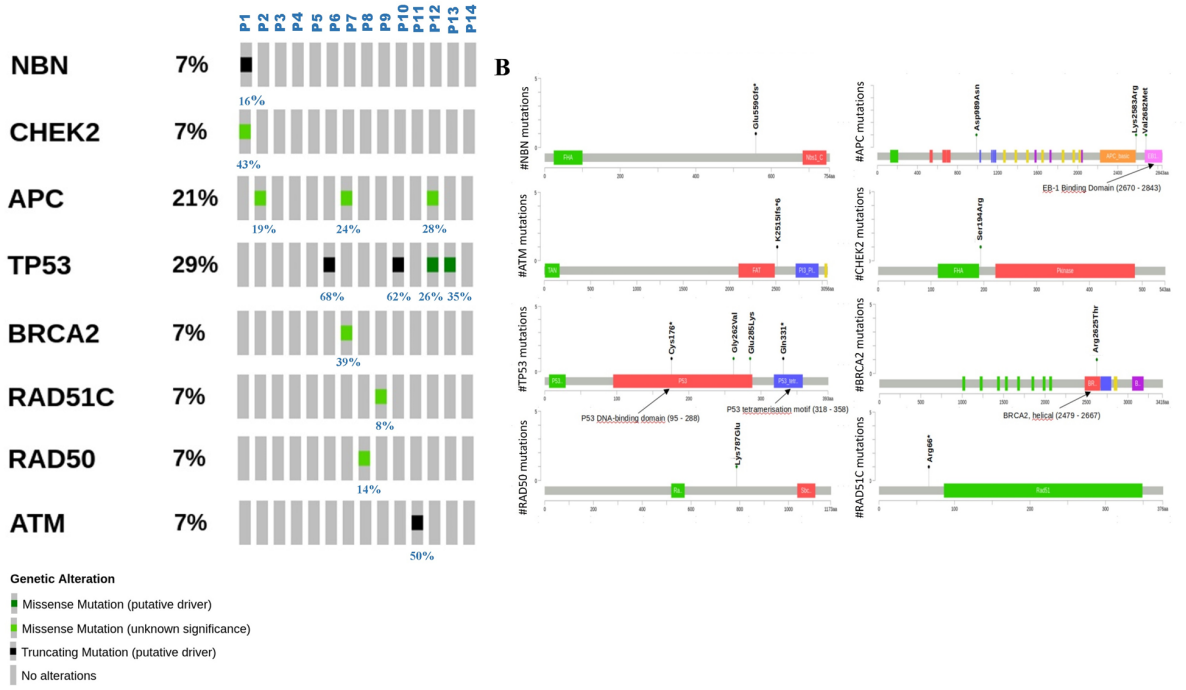

Figure 1. Mutational landscape of the 14 patients with BC. (A) OncoPrint showing the distribution and frequency of genomic alterations. The oncoprint was obtained using OncoPrinter (cBioPortal) for Cancer Genomic. Percentage in blue corresponds to variant allele frequency, and the letter P indicates patients' IDs. (B) Lollipop plots showing the distribution of identified mutations in BC patients on a linear representation of the proteins and their domains. The plots were obtained by the informatics tool Mutation Mapper-cBioPortal for Cancer Genomics. BC indicates breast cancer.

Table 2. Nonsynonymous SNPs predicted with SIFT, PolyPhen-2, MutationTaster, and LTRPred programs.

			MUTATION			PREDICTION				
PATIENT ID	GENE	GENE CHANGE	PROTEINE CHANGE	VARIANT ALLELE FREQUENCY(VAF) %	KNOWN MUTATION	SIFT	POLYPHEN-2	MUTATION TASTER	LTRPRED	TOTAL DELETERIOUS
P1	*CHEK2*	c.580A>C	p.S194R	43		T	B	B	N	0 of 4
P2	*APC*	c.2965G>A	p.D989N	19.63		D	D	D	D	4 of 4
P7	*APC*	c.7748A>G	p.K2583R	24		T	D	D	D	3 of 4
P7	*BRCA2*	c.7874G>C	p.R2625T	39.64	COSM7403374	D	D	D	D	4 of 4
P8	*RAD50*	c.2776A>G	p.K926E	14.33		D	D	B	D	3 of 4
P9	*RAD51C*	c.137T>A	p.L46H	8.51		D	D	D	D	4 of 4
P12	*APC*	c.8044G>A	p.V2682M	28.12	COSM4559410	D	D	D	D	4 of 4
P12	*TP53*	c.785G>T	p.G262V	25.71	COSM11198	D	D	D	D	4 of 4
P13	*TP53*	c.853G>A	p.E285K	35.48	rs112431538 COSM10722	D	D	D	D	4 of 4

Abbreviation: SIFT, Sorting Intolerant From Tolerant.

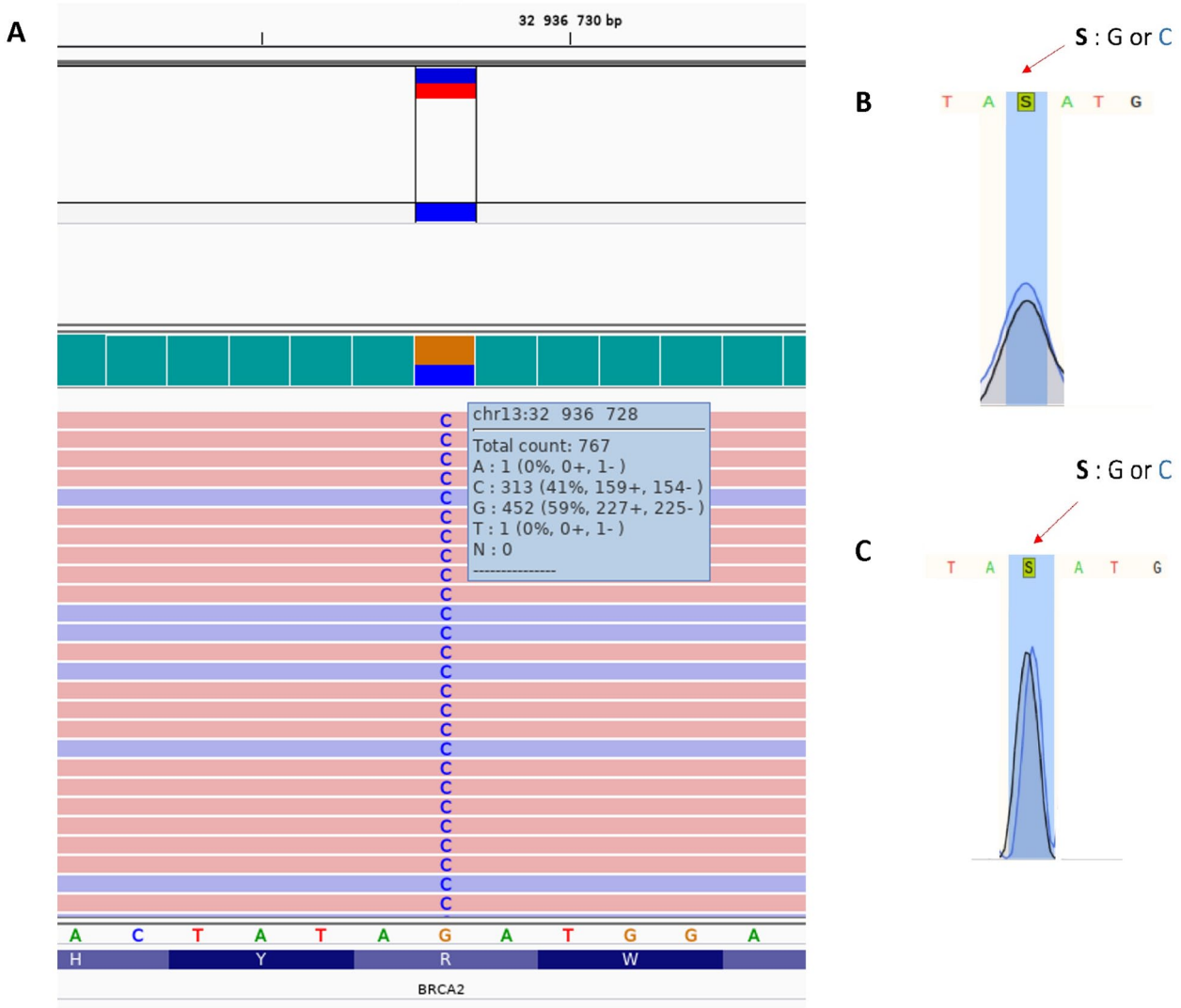

Figure 2. Integrative Genomics Viewer (IGV) and chromatogram visualization of c.7874G>C/p.R2526T mutation of BRCA2. (A) IGV screenshot of bam file analysis of a heterozygous substitution (c.7874G>C) of the BRCA2 gene. Among the total count of 767 reads in this position, there are 452 reads with G (reference allele) and 313 reads with C (alternative allele). This distribution showed that the 7874G>C mutation is a heterozygous. The confirmation of BRCA2 mutation c.7874G>C with Sanger sequencing in tumoral tissue sample (B) and in blood sample (C) revealed that the c.7874G>C is a heterozygous germinal mutation. BRCA2 indicates breast cancer type 2 susceptibility protein.

publication as a somatic bladder tumor mutation. We studied this mutation to see whether it was a somatic mutation of the breast tumor and also to see its impact on the protein. First, we validate the R2625T mutation using Sanger Sequencing, and then its impact on protein stability was assessed by molecular docking and MD.

The validation of the presence of R2625T using Sanger sequencing

To validate the presence of the R2625T mutation and to determine whether this mutation is somatic or germinal, we performed Sanger sequencing for tissue and blood DNA. The results showed that the mutation was present in both tissue and blood with 2 peaks at 7874 position of *BRCA2* gene (Figure 2B and 2C). The peaks correspond to G/C alleles; this was in agreement with the Illumina result where the depth of the mutated allele was 113/285 with a percentage of 40% (Figure 2A). The presence of 2 alleles in the Sanger and Illumina results indicates that this mutation is germinal and heterozygous.

The impact of R2625T mutation on the interaction between HD-OB1 and DSS1

The germinal R2625T mutation was predicted to be deleterious by all 4 functional prediction webservers. This mutation is located in the HD-OB1, a critical domain of BRCA2 through which it interacts with the DSS1 proteins. This led us to assume that the R2625T mutation may affect this interaction. In this regard, after the construction of the WT and MT 3D models of HD-OB1 in complex with DSS1 by molecular docking, we calculated their interaction free energy using the MM/GBSA method. To validate the used modeling/docking

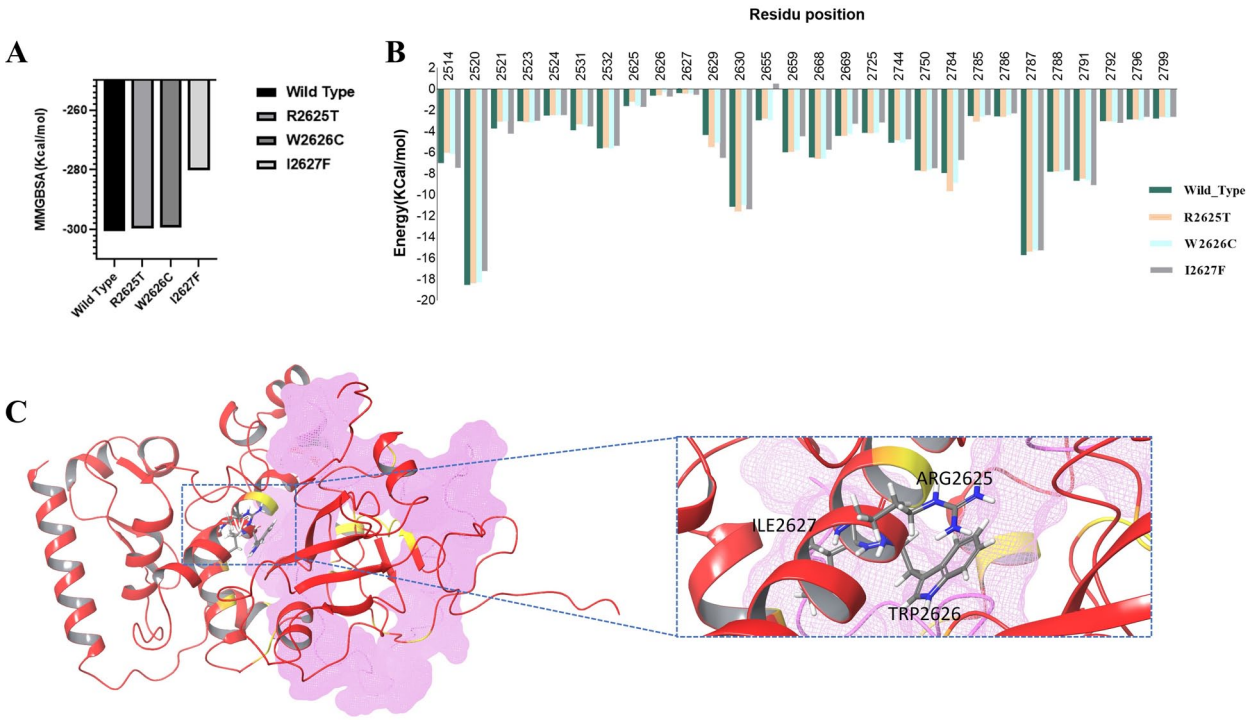

Figure 3. Analysis of the interaction between DSS1 and HD-OB1 in wild and mutated types (R2625T, W2626C, I2627F). (A) Binding free energy (MM/GBSA) of DSS1 in complex with HD-OB1. (B) Per residue free energy decomposition of the binding surface from HD-OB1-DSS1 complex (wild type and mutated type). Only the top 25 residues with high MM/GBSA score were presented. The positive and negative values indicate the unfavorable and favorable contribution for the binding, respectively. (C) Localization of R2625, W2626, and I2627 on the interaction surface of DSS1 and HD-OB1. The main top 25 residues implicated in the HD-OB1-DSS1 interaction are colored in yellow. HD indicates helical domain; MM/GBSA, molecular mechanics/generalized Born surface area.

protocol and the obtained results, 2 known pathogenic mutations (W2626I, I2627F) of BRCA2 HD-OB1were used as reference. The calculated free energy values obtained were –300.67, –299.79, –299.31, and –280.14 for WT, R2625T, W2626C, and I2627F, respectively (Figure 3A). These results revealed that the R2625T mutation does not affect the interaction between HD-OB1 and DSS1 as there is no significant difference in the calculated free energy value between the MT R2625T (–299.79 kcal/mol) and the WT model (–300.67 kcal/mol). The calculated free energy value of I2627F (–280.14 kcal/mol) showed a significant diminution compared to WT, while for the W2626C, the free energy value was approximately equal to that of WT. It was reported in a previous experimental study that the pathogenic W2626C mutation does not affect the binding affinity between BRCA2 and DSS1, while the I2627F mutation significantly decreases this affinity.[29] This was in agreement with our results and validates our docking/dynamic protocols. Furthermore, visual inspection of the interaction surface (Figure 3C) and calculation of the per residue MM/GBSA free energy decomposition showed that the R2625, W2626 and I2627 residue are not located in the main interaction surface with DSS1 and do not contribute to this interaction (Figure 3B). However, the interaction could be affected indirectly in the presence of 2627 mutation.

Indeed, a detailed observation of the top 25 residues of the protein-protein binding interface with the highest MM/GBSA scores (Figure 3B) showed a significant reduction in the MM/GBSA value at residue Q2655 in the mutated I2627F model; this value changed from a negative value (–2.96) in the WT to a positive value (+0.5) in the mutated I2627F model. This indicates that although residue I2627 does not have a direct impact on the interaction surface between *BRCA2* and DSS1, it indirectly influences this interaction through residue Q2655.

Impact of R2625T mutation on the dynamics of HD-OB1-DSS1 complex

The results of molecular docking-MM/GBSA were further validated under dynamic conditions through MD simulation. In this respect, we performed a comparative study of MD simulation between WT and 3 MT (R2625T, W2626C and I2627F) of HD-OB1 of BRCA2 in complex with DSS1. The MD results were analyzed using (1) protein RMSD that measures the conformational changes of given complex over time and describes whether the simulation is in equilibrium (Figure 4A) and (2) protein root mean square fluctuation (P-RMSF) that characterizes local changes along the protein chain (Figure 4B).[30,31]

A gradual increase in the RMSD fluctuations for all systems was observed during the initial 2 ns of the simulations. Subsequently, the systems started to stabilize around 4Å after

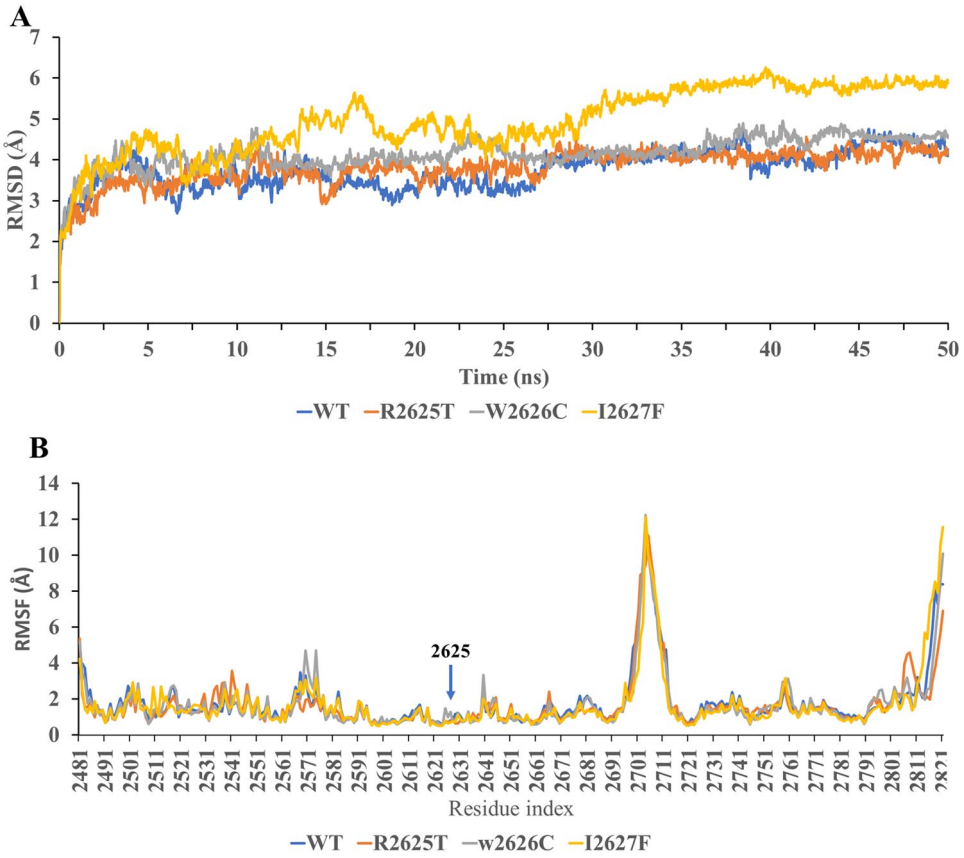

Figure 4. RMSD (A) and RMSF (B) analyses of C-alpha atoms for wild and mutated models (R2625T, W2626C, I2627F) of HD_OB1 in complex with DSS1 during 50 ns of MD simulations. The blue arrow indicates the position of residues 2625. HD indicates helical domain; MD, molecular dynamics; RMSD, root mean square deviation; RMSF, root mean square fluctuation; WT, wild type.

5 ns of the simulation, with WT, R2625T, and W2626C reaching thermodynamic equilibrium after 30 ns and remaining stable until the end of the simulations with average RMSD values of 3.96 ± 0.38Å, 3.81 ± 0.41Å, and 3.75 ± 0.47Å, respectively. This indicates that these mutations do not affect the stability of the HD-OB1-DSS1 complex as the RMSD results showed no significant difference between the R2625T, W2626C, and the WY, whereas, compared to WT, the RMSD of the I2627F system continued to increase slightly until 35 ns where it reached equilibrium and showed a higher RMSD value with an average of 4.96 ± 0.8Å, indicating that this mutation could affect the HD-OB1-DSS1 complex stability (Figure 4A). On the other hand, the RMSF values do not showed any significant differences between the 4 studied systems, indicating that all these mutations do not affect neither locally nor globally the protein flexibility (Figure 4B).

Taken together, the result of molecular docking-MM/GBSA and MD showed a significant concordance, indicating that the R2625T mutation has no potential impact neither on the interaction between HD-OB1 and DSS1 nor on the stability HD-OB1-DSS1 complex. However, we believe that further in vitro analyses are needed to validate these results.

Discussion

This study describes somatic/germinal mutations detected in Moroccan BC patients using a panel of 22 genes. We identified 11 SNPs and 2 indels in *TP53*, *APC*, *BRCA2*, *CHECK2*, *RAD50*, *RAD51C*, *NBN*, and *ATM* genes. Among these SNPs, 6 were predicted as deleterious by 4 software predictors. Among these 6 deleterious mutations, 4 were found in COSMIC database (TP53 c.853G>A(E285K) and c.853G>A(E285K) mutations, APC c.8044G>A(V2682M) mutation, and c.7874G>C(R2625T) *BRCA2* mutation). The TP53 c.853G>A mutation is a germinal deleterious mutation that has been reported in previous studies; it is located within the DBD of TP53 and it is a temperature-sensitive mutation inducing the loss of its function at 37°C.[32] E285K was shown to be associated with shorter survival in advanced *n*on–small cell lung cancer (NSCLC) treated with tyrosine kinase inhibitors (TKIs)[33]; furthermore, Blanden et al demonstrated in their study that E285K was the most unstable variant of TP53 and resistant to the zinc metallochaperone ZMC1.[34] In this study, we investigated one of the deleterious SNPs, c.7874G>C, that causes the change of arginine (R) to threonine (T) at position 2625 of the BRCA2 protein. This mutation has only been identified in a single study in a patient with bladder cancer and has been described/confirmed as a somatic mutation.[35] In contrast, in our study, we found that this mutation is germinal and heterozygous using Sanger sequencing for both tissue and blood DNA.

The R2625T mutation is located at the helical-binding domain of the DBD of BRCA2 protein, specifically at OB1 motif. It was demonstrated that mutations in this domain disrupt the interaction between BRCA2 and DSS1, a small acidic

protein required for the stability of BRCA2 which interrupts homologous DNA repair via BRCA2 in cells with DBD mutations, thus causing cancer evolution.[29,36] These findings led us to assume that the mutation R2625T might decrease DSS1-binding capacity with BRCA2 and inhibit its function. Molecular docking between DSS1 and WT/R2625T BRCA2 indicates that this mutation had no impact on HD-OB1-DSS1 interaction. Moreover, MD also showed that this mutation had no impact on HD-OB1-DSS1 complex stability. Indeed, the RMSD results showed no significant difference between the R2625T and the WT, with average RMSD values of 3.96 ± 0.38 and 3.75 ± 0.47, respectively.

Molecular docking and MD results were referenced to 2 known pathogenic mutations W2626C and I2627F. Experimental data from these mutations showed that W2626C does not affect or slightly affects the interaction between DSS1 and BRCA2, while I2627F highly affects this interaction.[29] Using our molecular docking and MD protocol, we found comparable results to those obtained from experimental data, which validates our docking and dynamics protocol and provides a baseline for comparison with R2625T results.

Although the results for R2625T showed that this mutation has no impact on the interaction between DSS1 and BRCA2, we believe that further experimental analysis is are needed, particularly by studying the impact of this mutation on the dislocation of BRCA2-DSS1 complex from nucleus to cytosol. Indeed, this phenomenon of dislocation alteration is observed especially in the presence of some mutations, that even retain residual DSS1-binding capacity have similar affect to that, ablate it more severely in impairing the nuclear localization of full-length BRCA2 protein.[29] For example, the W2626C mutation, which is adjacent to R2625T, retains significant DSS1 binding and also forms intracellular oligomers; however, its nuclear localization is impaired, which causes a lot of homologous DNA recombination.[29]

Collectively, this study provides an insight on the germinal/somatic mutational profile of Moroccan BC patients, which will help in the optimization of specific gene panel testing for Moroccan population. It is necessary to increase the number of the population size to have a global view. Furthermore, the in silico techniques used in this study could be used as proof-of-principle in the preliminary step for studying the impact of HD-OB1 mutations on HD-OB1-DSS1 complex stability.

Conclusion

This preliminary study aims to assess further potential germline /somatic mutations that may be related to BC using targeted gene panel sequencing. Collectively, the results of this study revealed the presence of 11 SNPs in various genes, such as *TP53, APC, BRCA2, CHECK2, RAD50, RAD51C*, *NBN*, and *ATM*. It is noteworthy that 6 of these SNPs were predicted to be deleterious. R2625T mutation was thoroughly investigated in this study as it is located in the BRCA2 HD-OB1 domain, a critical domain through which BRCA2 interacts with a stabilizing protein, the DSS1. Docking and MD demonstrated that this mutation had no impact on the affinity and stability of the HD-OB1-DSS1 complex. An in-depth experimental study is necessary to better understand the pathogenic mechanism of this mutation.

Acknowledgements

Financial support for this study was provided by the IRC (Institute of Cancer Research), the CNRST-PPR1 program, and the Biocodex Microbiota Foundation. The authors express their gratitude toward these funding sources.

Author Contributions

Conceptualization, S.K., E.M.B, O.Z and A.I.; methodology, S.K., E.M.B and O.Z.; software, S.K., E.M.B and O.Z.; validation, S.K., E.M.B, O.Z., A.I and B.E.; formal analysis S.K., E.M.B.; investigation, S.K., E.M.B., O.Z. A.I, and B.E.; resources, S.K., C.M, B.E.; data curation, S.K., E.M.B.; writing—original draft preparation, S.K., E.M.B and O.Z.; writing—review and editing, S.K., E.M.B., O.Z., R.E., A.I., L.B and B.E.; visualization, S.K., E.M.B and O.Z.; supervision, A.I, L.B, and B.E.; project administration, A.I, and B.E.; funding acquisition, A.I. Allauthors have read and agreed to the published version of the manuscript.

Data Availability Statement

Data supporting the findings of this study not presented within the article are available upon request.

Informed Consent

Approval for conducting this study was granted by the Ethics Committee of the Faculty of Medicine and Pharmacy in Rabat, Morocco (No. 103/17).

REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018;68:394-424.
2. Adeloye D, Sowunmi OY, Jacobs W, et al. Estimating the incidence of breast cancer in Africa: a systematic review and meta-analysis. *J Glob Health*. 2018;8:010419.
3. Kartti S, Bendani H, Boumajdi N, et al. Metagenomics analysis of breast microbiome highlights the abundance of Rothia genus in tumor tissues. *J Pers Med*. 2023;13:450.
4. Mavaddat N, Peock S, Frost D, et al. Cancer risks for BRCA1 and BRCA2 mutation carriers: results from prospective analysis of EMBRACE. *J Natl Cancer Inst*. 2013;105:812-822.
5. Shiovitz S, Korde LA. Genetics of breast cancer: a topic in evolution. *Ann Oncol*. 2015;26:1291-1299.
6. Slavin TP, Maxwell KN, Lilyquist J, et al. The contribution of pathogenic variants in breast cancer susceptibility genes to familial breast cancer risk. *NPJ Breast Cancer*. 2017;3:22.
7. Fong PC, Boss DS, Yap TA, et al. Inhibition of poly (ADP-ribose) polymerase in tumors from BRCA mutation carriers. *N Engl J Med*. 2009;361:123-134.
8. Seal S, Thompson D, Renwick A, et al. Truncating mutations in the Fanconi anemia J gene BRIP1 are low-penetrance breast cancer susceptibility alleles. *Nat Genet*. 2006;38:1239-1241.

9. Renwick A, Thompson D, Seal S, et al. ATM mutations that cause ataxia-telangiectasia are breast cancer susceptibility alleles. *Nat Genet*. 2006;38:873-875.
10. Rahman N, Seal S, Thompson D, et al. PALB2, which encodes a BRCA2-interacting protein, is a breast cancer susceptibility gene. *Nat Genet*. 2007;39:165-167.
11. Osborne C, Wilson P, Tripathy D. Oncogenes and tumor suppressor genes in breast cancer: potential diagnostic and therapeutic applications. *Oncologist*. 2004;9:361-377.
12. Singer CF, Balmaña J, Bürki N, et al. Genetic counselling and testing of susceptibility genes for therapeutic decision-making in breast cancer—An European consensus statement and expert recommendations. *Eur J Cancer*. 2019;106:54-60.
13. Fanale D, Incorvaia L, Filorizzo C, et al. Detection of germline mutations in a cohort of 139 patients with bilateral breast cancer by multi-gene panel testing: impact of pathogenic variants in other genes beyond BRCA1/2. *Cancers*. 2020;12:2415.
14. Zhang G, Zeng Y, Liu Z, Wei W. Significant association between Nijmegen breakage syndrome 1 657del5 polymorphism and breast cancer risk. *Tumour Biol*. 2013;34:2753-2757.
15. Maresca L, Lodovichi S, Lorenzoni A, et al. Functional interaction between BRCA1 and DNA repair in yeast may uncover a role of RAD50, RAD51, MRE11A, and MSH6 somatic variants in cancer development. *Front Genet*. 2018;9:397.
16. Win AK, Lindor NM, Jenkins MA. Risk of breast cancer in Lynch syndrome: a systematic review. *Breast Cancer Res*. 2013;15:R27.
17. Easton DF, Pharoah PDP, Antoniou AC, et al. Gene-panel sequencing and the prediction of breast-cancer risk. *N Engl J Med*. 2015;372:2243-2257.
18. Kilpivaara O, Vahteristo P, Falck J, et al. CHEK2 variant I157T may be associated with increased breast cancer risk. *Int J Cancer*. 2004;11:543-547.
19. Danecek P, Bonfield JK, Liddle J, et al. Twelve years of SAMtools and BCFtools. *Gigascience*. 2021;10:giab008.
20. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. 2010;38:e164.
21. Robinson JT, Thorvaldsdóttir H, Winckler W, et al. Integrative genomics viewer. *Nat Biotechnol*. 2011;29:24-26.
22. Robinson JT, Thorvaldsdóttir H, Wenger AM, Zehir A, Mesirov JP. Variant review with the integrative genomics viewer. *Cancer Res*. 2017;77:e31-e34.
23. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. *Nat Methods*. 2010;7:248-249.
24. Reva B, Antipin Y, Sander C. Predicting the functional impact of protein mutations: application to cancer genomics. *Nucleic Acids Res*. 2011;39:e118.
25. Drost H-G. LTRpred: _de novo_ annotation of intact retrotransposons. *J Open Source Softw*. 2020;5:2170.
26. Pettersen EF, Goddard TD, Huang CC, et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem*. 2004;25:1605-1612.
27. Yan Y, Tao H, He J, Huang SY. The HDOCK server for integrated protein–protein docking. *Nat Protoc*. 2020;15:1829-1852.
28. Weng G, Wang E, Wang Z, et al. HawkDock: a web server to predict and analyze the protein–protein complex based on computational docking and MM/GBSA. *Nucleic Acids Res*. 2019;47:W322-W330.
29. Lee M, Shorthouse D, Mahen R, Hall BA, Venkitaraman AR. Cancer-causing BRCA2 missense mutations disrupt an intracellular protein assembly mechanism to disable genome maintenance. *Nucleic Acids Res*. 2021;49:5588-5604.
30. Bouricha EM, Hakmi M, Kartti S, Zouaidia F, Ibrahimi A. Mechanistic evidence from classical molecular dynamics and metadynamics revealed the mechanism of resistance to 4-hydroxy tamoxifen in estrogen receptor alpha Y537S mutant. *Mol Simul*. 2022;48:1456-1463.
31. Bouricha EM, Hakmi M, Akachar J, Zouaidia F, Ibrahimi A. In-silico identification of potential inhibitors targeting the DNA binding domain of estrogen receptor α for the treatment of hormone therapy-resistant breast cancer. *J Biomol Struct Dyn*. 2022;40:5203-5210.
32. Lee DS, Yoon S-Y, Looi LM, et al. Comparable frequency of BRCA1, BRCA2 and TP53 germline mutations in a multi-ethnic Asian cohort suggests TP53 screening should be offered together with BRCA1/2 screening to early-onset breast cancer patients. *Breast Cancer Res*. 2012;14:R66.
33. Yang D, Han X, Li D, et al. Molecular diagnosis and clinical outcome of a lung cancer patient with TP53-E285K mutated Li-Fraumeni syndrome harboring a somatic EGFR-KDD mutation. *Am J Transl Res*. 2020;12:6689-6693.
34. Blanden AR, Yu X, Blayney AJ, et al. Zinc shapes the folding landscape of p53 and establishes a pathway for reactivating structurally diverse cancer mutants. *eLife*. 2020;9:e61487.
35. Hurst CD, Alder O, Platt FM, et al. Genomic subtypes of non-invasive bladder cancer with distinct metabolic profile and female gender bias in KDM6A mutation frequency. *Cancer Cell*. 2017;32:701-715.e7.
36. Li J, Zou C, Bai Y, Wazer DE, Band V, Gao Q. DSS1 is required for the stability of BRCA2. *Oncogene*. 2006;25:1186-1194.