Germ-Line Mutanome Profiling of The Breast Cancers in Pakistani Population

Sadia Ajaz¹, Sani-e-Zehra Zaidi¹, Saleema Mehboob Ali¹, Aisha Siddiqua², Muhammad Ali Memon²

¹ Dr. Panjwani Center for Molecular Medicine and Drug Research (PCMD), International Center for Chemical and Biological Sciences (ICCBS), University of Karachi, Karachi, Pakistan.
² Atomic Energy Medical Centre (AEMC), Jinnah Postgraduate Medical Centre (JPMC), Karachi, Pakistan.

Abstract

Statement of the Problem: Cancers are complex disorders. Consanguineous populations, by virtue of reduced genetic diversity, provide a unique model for the investigations of the underlying genetic component(s). A prototype example in the breast cancers is the identification of BRCA1/2 gene defects in Ashkenazi Jews. In Pakistan, the age-standardized rate (ASR) of breast cancer incidence in females is among the highest in Asia, whereas the mortality rate is one of the highest in the world. With the consanguinity rate of 56.4% and inbreeding coefficient (F) of 0.0331, it is extremely important to investigate the role of inherited mutations in breast cancers in Pakistani population.

Methodology and Theoretical Orientation: BROCA analysis for breast cancers consists of twenty-seven (27) established and candidate breast cancer genes involved in molecular carcinogenesis. A pilot hospital-based cohort study was designed. Eighty-five breast cancer patients and three controls with no medical history of any cancer, participated in the study. The BROCA investigations were carried out by a genomic capture, massively parallel next generation sequencing assay on Illumina HiSeq2000 assay with 100bp read lengths. Copy number variations were determined by partially-mapped read algorithm. Once the mutation was identified, it was validated by Sanger sequencing. After informed consent, the mutations were screened in the familial samples.

Findings: The analysis revealed germ-line mutations in 12% of the patients. These mutations were restricted to three genes (BRCA 1, BRCA 2, and TP53). The identified mutations consist of both novel and previously reported alterations and result in protein truncation. No mutations were identified in the remaining twenty-four (24) genes. Mutation screening in the familial samples identified carriers in four out of five families.

Conclusions and Significance: The study provides framework for the development of preventive and treatment strategies against breast cancers in Pakistani population.

Keywords: Breast Cancer, Susceptibility, Candidate Genes, Consanguinity, Pakistani Population.

Introduction

It is estimated that approximately 20-25% cases of breast cancers arise due to highly penetrant hereditary component. The information is mainly influenced by data from Caucasian populations. Pakistani population is highly consanguineous. The consanguinity rate is 56.4% [1] and inbreeding co-efficient is 0.0331 (Hussain and Bittles, 1998). Such populations provide a unique model to investigate and quantify the role of candidate genes and/or specific mutations in complex disorders. The hallmark example is the discovery of BRCA1&2 mutations in breast cancer patients among Ashkenazi Jews [2]. Information on genetic basis of breast cancers in Pakistani population is scarce. The region has one of the highest age-standardized incidence rates (ASIR) of breast cancers in Asia. Breast cancer-related mortality rates are amongst the highest in the world [3]. Despite the disease burden, little is known about the underlying factors, genetic and/or environmental, for breast cancers in this population.

The present pilot study is the first-systematic attempt to investigate the role of all known breast cancer genes in Pakistani population. We report the investigation of germline mutations in twenty-seven breast cancer candidate genes in a cohort of breast cancer patients from Pakistan. The study has significant implications in delineating the molecular pathology of inherited breast cancers.

Methods

Ethics Statement

The project is in accordance with the Declaration of Helsinki - 2013 [4]. Informed consent was obtained from the participants. The project was approved by Ethical Review Committees of collaborating institutions: Atomic Energy Medical Centre (AEMC), Jinnah Postgraduate Medical Centre (JPMC), and Dr. Panjwani Center for
Molecular Medicine and Drug Research (PCMD), International Center for Chemical and Biological Sciences (ICCBS), University of Karachi, Karachi, Pakistan.

Sampling

The pilot study, conducted from July 2016 to June 2017, included blood samples from ninety-four (94) breast cancer patients and two (02) healthy volunteers. 8-10ml of blood samples were transferred to the ACD-coated vacutainers (BectonDickinson®, Franklin Lakes, NJ, USA). In case of healthy volunteers, 5ml blood samples were collected. The samples were either processed immediately or stored at 4°C until DNA isolation. After quality and quantity control, eighty-five (85) samples were analysed by next generation sequencing (NGS).

Cases and Family History Instrumentation

Patients volunteered the information about the family history of breast cancer, or any other cancer, number of pregnancies, number of children, the age of menarche and age of menopause. Details on Bloom Richardson grading system and tumour node metastasis (TNM) staging as recommended by the Union International Contre le Cancer (UICC) were obtained from the hospital medical records with the patient’s consent.

DNA Extraction, Quantification, and Quality Control

Standard Phenol-Chloroform method was used for the extraction of DNA from the WBCs of patient’s blood sample [5]. Isolated DNA was quantified on UV spectrophotometer (Beckman Coulter® DU 530). The acceptable range of 260/280 ratios was set as 1.8-1.99. Quality control was carried out on 0.7% agarose gel and visualized under UV light using Azure C-300® imaging system. None of the extracted DNA samples showed the shredded or streaked pattern on the agarose gel (Figure 1).

Genomic Capture, Massively Parallel Next Generation Sequencing Assay (BROCA Analysis)

After DNA extraction, quantity and quality control, germline DNA was sequenced using breast and ovarian cancer gene panel (BROCA). This assay sequences all the exonic and flanking intronic sequences of the selected genes. The analysis is able to capture all types of mutations including large genomic re-arrangements in known breast cancer genes [6-8]. The breast cancer gene panel assay investigates following 27 genes: BRCA1, BRCA2, TP53, ATR, BARD1, BRIP1, FAM175A, FANCM, GEN1, MRE11A, NBN, RAD51B, RAD51C, RAD51D, RECQL, RINT1, SLX4, BAP1, PALB2, PTEN, STK11, XRCC2, ATM, CHEK1, CHEK2, CDH1, and CTNNA1.

Exon Amplification and Sanger Sequencing

The exons in BRCA1, BRCA2, and TP53 genes carrying mutations as identified by N.G.S. were amplified. The amplicons were purified by an isopropanol-based method. 15µl of the purified amplified products along with the primers were used for sequencing. Sanger sequencing of samples was performed commercially.

Bioinformatic Analysis

ExAC and NHGRI databases were used as reference for the identification of novel mutation(s).

Results and Discussion

10 samples from 85 patients (approx.12%) had a germ-line mutation. These were confined to BRCA1, BRCA2, and TP53 genes.

BRCA1

In total five mutations were identified in BRCA1 in case series of 85 patients. Among these two are novel. Details of a representative mutation are provided below:

c.68_69del (Exon 2) Mutation: The mutation started at 17:41276045 (hg37). It was validated by Sanger sequencing (Figure 2A and 2B). The consequence of mutation is 23frameshift leading to stop codon at amino acid position 39 and a truncated protein.

Age of the Proband and Familial Segregation Analysis: Proband is a 43 years old female. After the parental consent, c.68_69del germline mutation analysis in DNA samples from daughters of the patient revealed the absence of mutation (Figure 3).

Comparison of Identified Mutation with Published Reports from Different Populations: The mutation along with two other founder mutations (BRCA1 5382insC and BRCA2 6174delT) accounts for 60% of hereditary cases of breast cancers in Ashkenazi Jews [9]. It contributes to 16%-20% of breast cancer cases diagnosed before the age of 50 years in this population. The mutation has 2068 BIC database entries. It has been identified in American, Ashkenazi Jewish, Ashkenazi/Central European, Austrian, Belgian, British, Canadian, Central/Eastern Europe, Chilean, Czech, Dutch, French, German, Hispanic, Indian, Italian, Latvian, Malaysian, Mexican American, Native American, Norwegian, Spanish, USA, Western Europe, Western Europe/Ashkenazi populations (NCBI).

This is the sixth report of c.68_69 (185delAG) mutation from Pakistan (6/686: 0.87%) and the first report of this mutation in an Urdu-Speaking ethnic group. The mutation has previously been identified in patients belonging to Punjabi and Pathan ethnicity. The geographical location of reported mutations is shown in the map of Pakistan (Figure 4).

Figure 1: DNA quality control on 0.7% agarose gel.

Figure 2: BRCA1 c.68_69del (185delAG) mutation detection by Sanger sequencing. Fig. 2.A. Control sample. Fig. 2.B. Patient’s sample.

Figure 3: Pedigree of breast cancer patient carrying c.68_69del (185delAG) mutation. V: Variant sequence, N: Wild-type (Normal) sequence.
Discussion

**BRCA1** mutations play a significant role in genetic predisposition to cancers. It consists of 24 exons. The mutations are more frequent in exons encoding three specific domains of BRCA1 protein (Figure 5). The first is RING (Really Interesting New Gene) domain encoded by exons 2-7. The second domain is encoded by exons 11-13, which comprises 65% of the BRCA1 protein sequence. It includes two Nuclear Localization Signals (NLS), and a serine-rich component. It is also the region where key DNA repair proteins, RAD50 and RAD51, as well as c-Myc (a transcription factor) and Retinoblastoma (a cell cycle regulator) binds. The third domain is BRCA1 C-terminal (BRCT) domain. It recognizes pSer-X-X-Phe sequence and forms a complex with its phosphorylated binding partners. These include BACH1, CtIP, and CCDC98/abraxas [10].

The RING domain of the BRCA1 plays a key role in its association with BARD1 (Figure 6A), with consequent ubiquitin ligase activity and nuclear retention of two proteins. c.68_69del (185delAG) results in Glu23fs and stop codon at position 39. The transcript does not undergo non-sense mediated mRNA decay (NMD) and results in a truncated protein (Figure 6B). It is known to affect apoptotic pathway, chemosensitivity, and transcription/gene regulation. In specific populations, it is the most frequently reported founder mutation [11].

The present study reports c.68_69 as candidate BRCA1 germline founder mutation, underlying predisposition to breast cancers in the local population. Despite the limited sample size in previous and present reports, this germline mutation has been identified in two patients from Punjabi ethnicity, two from Pathan ethnicity, and one where ethnicity is unknown. The present study reports this mutation in a patient belonging to Urdu-speaking ethnicity, with a strong family history of multiple primary tumors including five cases of breast cancers. These recurrent reports indicate that the mutation is a strong candidate founder mutation in this population.

Taking into account the present study as well as previous reports [12,13], the frequency of this mutation in breast cancer patients from Pakistani population is close to that of Ashkenazi Jews. It is possible that the mutation may have arisen prior to separation of the Jewish diaspora, about 2000 years ago [14]. There is also some evidence that the mutation may have arisen twice, once in Jewish and once in non-Jewish population. However, the data is inconclusive. It has been proposed, that the reason for its persistence in different populations is possibly as yet unidentified biological advantage [15]. The evolutionary insight into the mutation is as yet an intriguing conundrum, which might hold the potential in translational medicine.

In conclusion, the present study provides evidence for a founder mutation in Pakistani population. The determination of recurrent mutation frequency has important implications in devising strategies for genetic testing, counselling, and treatment of familial/hereditary breast cancers in the indigenous population.

**BRCA2**

In total three **BRCA2** mutations were identified in a case series of 85 patients. Details of a representative mutation are provided below:

**c.5642delAATC** (Exon 11) Mutation: The mutation started at 13: 32914134 (hg37). It was validated by Sanger sequencing. The consequence of mutation is 1881frameshift leading to stop codon at amino acid position 1891 and a truncated protein (Figure 7).

Analysis of Germline Mutations: Comparison with NHGRI database shows that although 1881fs has been reported earlier, underlying germline mutation c.5642delAATC is being reported for the first time.
BRC repeats (Figure 9). It is through this domain that BRCA2 forms a complex with Rad51. Thus, through interaction with this recombinase, the role of BRCA2 in double strand break (DSB) repair was established. The formation of BRCA2-Rad51 complex involves two domains of BRC repeats, the other one at the C-terminus [6]. The binding at BRCA2 C-terminus leads to oligomerization of Rad51. During homologous DNA repair (HDR), DSBs are first cleaved and converted into single stranded breaks (ssDNA) coated by Rad51 nucleoprotein filaments, which facilitate the invasion of ssDNA into the homologous sister strand. The invading ssDNA serves as a primer for DNA synthesis, using the undamaged sister strand as the template, enabling the strand exchange. Thus, HDR is error-free [16]. Mutation in BRCA2 exon 11, is therefore likely to result in the accumulation of DSBs and genome instability.

TP53:
In all, two mutations in TP53 gene were identified in a case series of 85 patients. Details of a representative mutation are provided below:
Discussion: TP53 is known to be the guardian of genome [17]. It plays a central role in multiple primary cancers. Although a high number of somatic mutations have been reported, the germline mutations are comparatively rare [18]. As evidenced by BRCA1 and BRCA2 genes in Ashkenazi Jews, consanguineous populations provide a unique model for the identification and quantification of role of specific genetic alterations in cancer pathology. A previous study from different region of the country investigated 105 familial/early onset cases of breast cancer for germline mutations in TP53. However, they did not find any nonsense mutation in their series [19]. The differences in population sub-stratification and the methodology must be taken into consideration while comparing these two studies. In the present study the investigations were carried out by a genomic capture, next generation sequencing assay with 100bp read lengths (BROCA).

W91X mutation arising from G>A substitution in exon 4 was identified in a 48-year-old Gujrati patient, who reported strong family history of breast cancers. This is the first time this germline mutation is being reported. Recent data suggests that TP53 truncating mutations are likely to have neomorphic potential.

Interestingly, at least some of the truncated p53 mutants are able to evade nonsense-mediated decay, resulting in promotion of metastasis and tumour maintaintance similar to gain-of-function mutants [20]. Figure 12.

Given the emerging data regarding the functional role(s) of truncated p53 and present report of one such mutations, TP53 may be considered in familial cases of breast cancers [21-29]. The present study has shown that genetic component other than BRCA1 and BRCA2 plays a definitive role in susceptibility to breast cancers in the local population [30-37].

In conclusion, the present report is the first comprehensive study of twenty-seven breast cancer genes in representative cases from Pakistan [37-42]. The discovery study shows that mutations in one of three genes, BRCA1, BRCA2, and TP53 comprise the germline mutanome profile of breast cancers in Pakistani population [43-48].

Additional Information

Authors’ contributions: SA-Study design, Sampling, Benchwork, Analysis, and Manuscript Preparation; SZZ - Benchwork, Analysis; SMA- Sampling, Benchwork, Analysis; AS – Sampling, Medical Data Collection, MAM - Sampling, Medical Data Collection;

Acknowledgments:

The research was funded in part by the ICCCBS recurrent grant. Prof. MC King and Dr. Tom Walsh (Department of Genome Sciences and Medical Genetics, University of Washington, Seattle, Washington, USA) were instrumental in BROCA analysis and its sponsoring. The staff at AEMC, JPMMC, Karachi for their facilitation, and transport office at the ICCCBS for sampling. Finally, we are especially grateful to the patients and their families who participated in the study.

Figure 12: Locus and structure-function correlation of TP53 gene and protein. c.272G>A (W91X) in exon 4 leads to protein truncation.

Conflicts of Interest:

The authors declare no potential conflicts of interest.

Financial Statement:

The funders had no role in study design, data collection, analysis, and decision to publish or preparation of the manuscript.

All procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1975, as revised in 2013 [4]. Informed consent was obtained from all patients and participants included in the study.

References


23. Breast Cancer Information Core (BIC) database


37. National Center for Biotechnology Information


41. PED 6.0.2 – Pedigree Drawing Software


43. PyMol: https://pymol.org/2/


