Clinical characteristics and exploratory genomic analyses of germline BRCA1 or BRCA2 mutations in breast cancer

Authors:
Sehhoon Park1,*, Eunjin Lee2,*, Seri Park3, Sohee Lee3, Seok Jin Nam4, Seok Won Kim4, Jeong Eon Lee4, Jong-Han Yu4, Ji-Yeon Kim1, Jin Seok Ahn1, Young-Hyuck Im1, Woong-Yang Park2, Kyunghee Park2,§, Yeon Hee Park1,3,§

Author Affiliations:
1Division of Hematology-Oncology, Department of Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Korea
2Samsung Genome Institute, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Korea
3Samsung Biomedical Research Institute, Samsung Medical Center, Sungkyunkwan University, School of Medicine, Seoul, Korea
4Department of Surgery, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Korea

*These authors contributed equally.
§Co-Corresponding Author:
Yeon Hee Park, MD, Ph.D.: E-mail: yhparkhmo@skku.edu
Division of Hematology-Oncology, Department of Medicine, Samsung Medical Center
Abstract

gBRCA1/2 mutations increase the incidence of breast cancer (BC) by interrupting the homologous recombination repair (HRR) pathway. Although gBRCA1 and gBRCA2 BC have similar clinical profiles, different molecular characteristics have been observed. In this study, we conducted comprehensive genomic analyses and compared gBRCA1/2 BC. Sanger sequencing to identify gBRCA1/2 mutations was conducted in 2,720 patients, and gBRCA1 ($n=128$) and gBRCA2 ($n=126$) mutations were analyzed. Within that population, deep target sequencing (TS) and matched whole transcriptome sequencing (WTS) results were available for 46 and 34 patients, respectively. An internal database of breast-cancer patients with wild-type gBRCA was used to compile a TS ($n=195$) and WTS ($n=137$) reference dataset.

Three specific mutation sites, p.Y130X ($n=14$) and p.1210Afs ($n=13$) in gBRCA1 and p.R294X ($n=22$) in gBRCA2, were comparably frequent. Immunohistochemistry subtyping determined that the incidence of triple negative BC was higher among those with a gBRCA1 mutation (71.9%), and estrogen receptor (ER)-positive BC was dominant in those with a gBRCA2 mutation (76.2%). gBRCA1/2 mutations were mutually exclusive with PIK3CA somatic mutations ($P<0.05$), and gBRCA1 frequently co-occurred with TP53 somatic mutations ($P<0.05$). The median tumor mutation burden was 6.53 per megabase (MB) in gBRCA1 and 6.44 per MB in gBRCA2. The expression of AR, ESR1, and PGR was significantly upregulated with gBRCA2 mutation compared with gBRCA1 mutation.

gBRCA1 and gBRCA2 BC have similar clinical characteristics, but they have different molecular subtypes, co-altered somatic mutations, and gene expression patterns.

Implications: Even though gBRCA1 and gBRCA2 mutations both alter HRR pathways, our results suggest that they generate different molecular characteristics and different mechanisms of carcinogenesis.
Among Asian breast cancer patients, 2 to 3% are reported to harbor mutations in germline BRCA1 or BRCA2 (1-4). Compared with the general population, the incidence of breast cancer in these patients is increased by 69 to 72% until the age of 80 years. Germline mutations in both BRCA1 and BRCA2 also increase the incidence of breast and many other cancers, such as ovarian, peritoneal, pancreatic, and prostate (5, 6). These risks are understood to be determined by the position and type of mutation in BRCA1/2, as assessed by the Consortium of Investigators of Modifiers of BRCA1/2 (7).

Both BRCA1 and BRCA2 function in DNA damage repair (DDR) by playing a major role in the homologous recombination repair process. The BRCA1 protein induces a 5′ to 3′ resection of double-strand breaks to generate overhanging 3′ single-strand DNA (ssDNA) and localizes damaged DNA by loading DNA recombinase RAD51 onto ssDNA (8).

gBRCA1- and gBRCA2-mutant breast cancer have similar clinical profiles but different molecular characteristics. The onset of the disease differs between them in that the incidence increases until age 30 to 40 years for gBRCA1 and age 40 to 50 years for gBRCA2, though it then remains constant until age 80 in both cases (9). The proportion of molecular subtypes also differs in that gBRCA1-mutated breast cancer is predominantly triple negative breast cancer (TNBC), whereas gBRCA2-mutated breast cancer is predominantly estrogen receptor (ER)-positive (10, 11). Nevertheless, both gBRCA1- and gBRCA2-mutant breast cancer show superior progression-free survival (PFS) and a higher overall response rate with a poly adenosine diphosphate-ribose polymerase inhibitor than with standard treatment, which indicates that gBRCA1/2 mutant breast cancer is highly dependent on other DDR pathways regardless of its underlying molecular subtype (12, 13).

Based on the unique characteristics of gBRCA1/2 mutations in breast cancer, we
performed comprehensive clinical and genomic analyses of samples from operable breast
cancer patients with gBRCA1/2 mutations.

Patients and Methods

Study population and data collection

The medical records of patients who received a germline BRCA1/2 test at Samsung
Medical Center (SMC) between January 2007 and October 2018 were retrospectively
reviewed (n = 2,720). The study population was selected using the following criteria: patients
with pathologic gBRCA1 or gBRCA2 variants who received surgery with curative intent,
were diagnosed at least one year before data acquisition, and had adequate clinical
information available for analysis (Fig. 1A). Pathology and immunohistochemistry (IHC)
data were collected from surgical and biopsy reports. Genomic analyses were performed on
samples with available genomic outcomes using a deep target sequencing (n = 45) protocol
called CancerSCAN™ and whole transcription sequencing (WTS, n = 34). The patients with
WTS data were a subset of the population with deep target sequencing outcomes. Genomic
data from an internal database (n = 195) were used as a gBRCA wild type (WT) reference
(Fig. 1B). This study was approved, and written informed consent was waived by the SMC
Institutional Review Board (IRB#2019-03-067). The data that support the findings of this
study are available from the corresponding author upon reasonable request.

Germline BRCA mutation test

Genomic DNA was isolated from peripheral blood leukocytes. Germline BRCA1/2
mutation status was tested by direct Sanger sequencing. Annotation used a pre-defined
internal calling algorithm based on previous reports. Only patients with a pathogenic variant
in the reports were considered to have a gBRCA mutation. Patients with equivocal variants and variants of unknown significance were excluded from the analyses.

**Target sequencing (CancerSCAN™) and whole transcriptome sequencing**

Target sequencing results were extracted from previously calculated CancerSCAN™ data. CancerSCAN™ is designed to enrich the exons of 381 genes curated from the literature (Supplementary Table S1). DNA (250 ng) from cancer tissue was sheared in a Covaris S220 ultrasonicator (Covaris, Woburn, MA, USA) and used to construct a library with CancerSCAN™ probes and a SureSelect XT reagent kit, HSQ (Agilent Technologies) according to the manufacturer’s protocol (14, 15). After being enriched, exome libraries were multiplexed and sequenced using the 100-bp paired-end mode in a TruSeq Rapid PE Cluster Kit and TruSeq Rapid SBS kit on the Illumina HiSeq 2500 sequencing platform (Illumina Inc., San Diego, CA, USA). The DNA sequence data were aligned to the human genome reference (hg19) using the MEM algorithm in BWA 0.7.5 (16). Duplicate read removal was performed using Picard v.193 and SAMTOOLS v0.1.18 (samtools.sourceforge.net). Local alignment was optimized using the Genome Analysis Toolkit (GATK) v3.1-1 (https://software.broadinstitute.org/gatk/). BaseRecalibrator from GATK was used for the recalibration based on known single nucleotide polymorphisms (SNPs) and insertion/deletions (indels) from Mills, dbSNP138, 1000G gold standard, 1000G phase1, and Omni 2.5. Single nucleotide variations (SNVs) were detected using MuTect and LoFreq (17, 18). Falsely detected variants were filtered out using a script developed in-house. ANNOVAR was used to annotate the detected variants with dbSNP138, the Catalogue of Somatic Mutations in Cancer (COSMIC), TCGA, and an in-house Korean SNP database. Indels were detected by Pindel (19). Germline variants were filtered out by removing the
variants with an allele frequency greater than or equal to 97%, and suspected germline variants were removed based on an allele frequency greater than or equal to normal Korean samples. Loss of heterozygosity (LOH) status is calculated using PureCN (R-package) (Supplementary Fig. S1A, Supplementary Table 11).

For WTS, sequencing libraries were prepared using the TruSeq RNA Sample Preparation kit v2 (Illumina). Sequencing of the RNA libraries was performed on an Illumina HiSeq2500 in the 100-bp paired-end mode of a TruSeq Rapid PE Cluster kit and TruSeq Rapid SBS kit. After trimming poor-quality bases from the FASTQ files obtained from the sequencing, we aligned the reads to the human reference genome (hg19) with TopHat (v 2.0.6) (20) and performed a reference-guided assembly of transcripts with Cufflinks (v 2.1.1) (21). Alignment quality was verified with SAMTOOLS (v 0.1.19) (22). Gene expression was estimated from the RNA-Seq data of the patients using a count-based method with HTSeq (22). We selected 20,345 protein-coding genes, and genes expressed in at least three samples were retained. A total of 16,971 genes were thus considered for analysis. Gene counts were used as the input for trimmed mean of M value normalization by the R package edgeR (23), and normalized counts were transformed to log2-counts per million by applying voom from the R package limma (24) (Supplementary Fig. S1B). Raw data is available at NCBI SRA (PRJNA625821) and GEO (GSE149276).

Tumor mutation burden, PAM50 calculation, and gene set analyses

Tumor mutation burden (TMB) was defined as the number of somatic variants per megabase (MB) of the genome. We assessed TMB based on variants detected by the targeted sequencing in CancerScan™. Among the variants in the coding region, nonsynonymous SNVs and frameshift indels were counted. Germline variants with a population frequency >
0.001 in the Exome Aggregation Consortium (25) database and the Korean population were excluded. Variants listed in COSMIC (26) were filtered. To calculate the TMB per MB, the total number of mutations counted was divided by the size of the coding region in the targeted region.

We performed intrinsic subtyping with log-scaled normalized expression values using the 50-gene Prediction Analysis of Microarray (PAM50) subtype predictor, as described by Parker et al. (27). The PAM50 subtype predictor classified tumors into the following groups: Luminal A, Luminal B, HER2-enriched, basal-like, and normal-like.

Gene set analyses were performed using R package GSVA (28), which calculates the gene set enrichment score for each sample based on gene expression data. For the gene set enrichment analysis, we used the Molecular Signatures Database (MSigDB v6.2, http://software.broadinstitute.org/gsea/msigdb/index.jsp), which is a collection of 17,810 annotated gene sets, including 50 hallmark gene sets (29).

**Statistical methods**

Descriptive statistics are used to describe the characteristics of the study population. Overall survival (OS) and PFS in patients with gBRCA1-mutated versus gBRCA2-mutated status and gBRCA1/2-mutated versus WT status were analyzed using the Kaplan-Meier method.

The Maftools R package (30) was used to present the somatic mutation profile and detect a mutually exclusive set of genes, which were identified by Fisher’s exact test. To compare the TMB between gBRCA1/2 and WT, ANOVA and t-testing were used. Fisher’s exact test was used to identify genes whose mutations were associated with the TN or ER+ subtypes.

Differentially expressed gene (DEG) analyses for gBRCA1/2 status used a two-group t-test.
The raw p-value was adjusted to a false discovery rate (FDR). Genes within the cut-off criteria of FDR $P < 0.01$ and an absolute fold change $>2$ were designated as DEGs for gBRCA1/2. The thresholds of $P < 0.05$ and absolute fold change $>1.5$ were used for significant DEGs between gBRCA1/2 and WT. The thresholds for gene set enrichment analyses between gBRCA1 and gBRCA2 were $P < 0.01$ and an absolute difference in GSVA score $>0.3$. For GSVA between gBRCA1/2 and WT, both $P < 0.01$ and an absolute difference in GSVA score $>0.2$ were required.

A significance level of $P < 0.05$ was generally held to be significant in the analyses. All statistical tests were performed using R software v.3.4.2 (https://www.r-project.org/).

Results

Baseline demographics of the study population

Descriptions of the baseline demographics are given by gBRCA1 ($n = 128$) and gBRCA2 ($n = 126$) mutation status (Fig. 1A). The median age was 39 years old (range 25–67) for patients with gBRCA1 mutations and 40 years old (range 23–68) for patients with gBRCA2 mutations. Most patients were tested for the gBRCA mutation due to a family history of either breast or ovarian cancer (65.7%), followed by bilateral breast cancer (21.3%), and disease diagnosed at younger than 40 years old (18.0%). Most patients were in a pre-menopausal state (81.2%), and this ratio was similar in patients with gBRCA1 (89.1%) and gBRCA2 (73.0%) mutations. Patients received either modified radical mastectomy and total mastectomy (37.5%) or breast-conserving surgery and partial mastectomy (62.6%). Clinical stage at diagnosis was stage III in 24.5%, stage II in 40.6%, and stage I in 30.3%, according to the American Joint Committee on Cancer, Breast Cancer Staging, 7th edition. The main histopathology type was invasive ductal carcinoma (86.2%) for both gBRCA1 mutations.
(86.7%) and gBRCA2 mutations (85.7%). Details of the baseline demographics are provided in Table 1.

Clinical outcomes of the study population

Median follow-up duration of the study population was 55.3 months (range 13.7–261.7). Neoadjuvant chemotherapy was applied in 24.8% of patients, and 77.6% of patients received adjuvant treatment. Disease recurrence after surgery was observed in 26.4% of the study population, and the median time to recurrence was 147.3 months (range 127.4–186.2) for patients with a gBRCA1 mutation and 114.8 months (range 95.2–183.6) for those with a gBRCA2 mutation. Palliative chemotherapy was applied in 9.5% of patients, and 5.5% of patients had died by the time of data cut-off (Table 2). Survival analyses between gBRCA1 and gBRCA2 showed no significant difference in OS ($P = 0.21$), recurrence-free survival (RFS, $P = 0.98$), or time to distant metastases ($P = 0.82$) (Fig. 1C). Post hoc analyses conducted with individual loci showed that patients with mutations in gBRCA2 p. Lys467X had comparably shorter RFS ($P = 0.05$) than those with other gBRCA2 mutations. In a similar manner, RFS compared in a range of 200 amino acids showed that patients with BRCA2 mutations in 400 to 600 ($P < 0.01$) and 800 to 1,000 ($P = 0.02$) amino acids had shorter RFS than the rest of the patients (Supplementary Fig. S2). No specific locus in BRCA1 correlated with RFS. In addition, survival analyses conducted based on ER status dose not showed difference between gBRCA1 and gBRCA2 mutation. As an exploratory analysis, we looked in to the pathologic complete response (pCR) rate based on gBRCA mutation. In gBRCA1 mutant, 14 patients among 35 patients (40.0%) who received neoadjuvant chemotherapy achieved pCR. However, gBRCA2 mutant only showed 1 pCR among 28 patients (3.6%). Similarly, in wild type patients, we observed pCR in 8% of the
patients.

Germline BRCA1/2 mutation profiles and co-altered somatic mutations identified by CancerSCAN™

The mutations that occur in gBRCA are mostly frameshift mutations (45.2%), followed by nonsense mutations (43.3%), splicing mutations (7.1%), and a missense mutation (2.0%) (Supplementary Table S2). Among the mutations, Tyr130X (10.9%) and Glu1210Argfs (10.2%) in gBRCA1 and Arg2494X (17.5%) in a helical domain of gBRCA2 were observed frequently. Other mutations were observed in a broad range of loci not limited to a specific domain (Fig. 1D). We assessed co-altered somatic mutation profiles in samples available for target sequencing (n = 45, gBRCA1 n = 25, gBRCA2 n = 20) and identified co-altered somatic mutations in 44 (97.8%) of the samples. Mutations in TP53 (60%), FAT3 (33%), and FANCI (27%) were commonly observed. The mutations identified in BRCA1/2 were identical to previously detected germline mutations (Fig. 2A). A test for co-occurrence and exclusivity shows that gBRCA2 mutations were mutually exclusive with somatic mutations in TP53 (35.0% P < 0.05) and MET (0%, P < 0.05). In contrast, gBRCA1 mutations highly co-occurred with somatic mutations in TP53 (84.0%, P < 0.05) and MET (24.0%, P < 0.05) (Fig. 2B).

To identify the overall difference in somatic mutations between gBRCA mutants and the WT, target sequencing results from the merged gBRCA1/2 samples (n = 45) were compared with the results from the gBRCA WT reference samples (n = 195). Interestingly, the PIK3CA mutation was mutually exclusive with gBRCA mutation (P < 0.05). Specifically, only 13.3% of patients with a gBRCA mutation had a somatic alteration in PIK3CA. However, the somatic PIK3CA mutation rate was 27.7% in patients with WT gBRCA (Fig. 2C and...
Supplementary Fig. S3).

Median TMB was 6.525 per MB in gBRCA1, 6.437 per MB in gBRCA2, and 6.439 per MB in the WT reference patients. There were no significant differences between the groups (Fig. 2D).

**Immunohistochemistry subtypes and PAM50 molecular subtypes**

In the IHC subtyping, patients with gBRCA1 mutations were more frequently diagnosed with TNBC (71.9%), and those with gBRCA2 mutations were more frequently diagnosed with ER-positive breast cancer (76.2%). This result was compared with the PAM50 results from those available for WTS (n = 34) (Fig. 3A). Among the TNBC patients diagnosed by IHC, 87% were still predicted as a basal type by PAM50. A significant discrepancy was also observed in the ER-positive subjects analyzed by IHC: 31% of them (n = 5) were predicted as a basal subtype by PAM50. These results were compared with the expression profile of ESR1, which was arbitrarily classified into high (n = 13), intermediate (n = 8), and low (n = 13). As expected, in most of the patients with a gBRCA2 mutation (73%), ESR1 was upregulated, but ESR1 expression was either intermediate or low (90%) in those with a gBRCA1 mutation (Fig. 3B). In terms of the gene expression pattern of representative genes, the samples rated ER-positive by IHC and basal type by PAM50 had relatively upregulated BRCA2 expression (Supplementary Fig. S4).

Based on the IHC molecular subtype, we again compared the somatic mutation profiles with those of patients with WT gBRCA. A comparison between TNBC patients with a gBRCA1 mutation or WT found a higher ARID1B somatic mutation rate in patients with WT gBRCA, whereas the patients with a gBRCA1 mutation had a higher RET and PIK3CA somatic mutation rate (Fig. 3C). Similarly, the comparison between ER-positive patients with
a gBRCA2 mutation and WT gBRCA showed a higher PIK3CA somatic mutation rate in patients with WT gBRCA, whereas patients with a gBRCA2 mutation had a higher NOTCH1 and ZNF217 somatic mutation rate (Fig. 3D).

**Gene expression comparison between gBRCA1 and gBRCA2**

Initial analyses were conducted by comparing the gene expression patterns of patients with a gBRCA1 mutation (n = 19) with those with a gBRCA2 mutation (n = 15) (Fig. 4A).

Notably, the gBRCA2-mutated samples showed upregulation in hormone receptor or hormone-induced genes, such as AR (FDR=0.001), KIAA1324 (FDR=0.003), and HSPB8 (FDR=0.004); membrane tyrosine kinases, such as ERBB4 (FDR=0.002); and other breast cancer–related genes, such as BCAS1 (FDR=0.005) and GATA3 (FDR=0.006). On the other hand, the gBRCA1 samples showed upregulation in cytokines, such as CXCL5 (FDR = 0.008) (Fig. 4B, Supplementary Fig. S5, and Supplementary Table S3). Although it did not satisfy the pre-defined threshold, well-known breast cancer genes such as ERBB3, ESR1, PGR, and BRCA1 were upregulated in gBRCA2 samples, and EGFR and PIK3CA were upregulated in gBRCA1 samples (Fig. 4C).

The top-ranked gene set identified by GSVA (FDR < 0.002) in the gBRCA1 mutants was consistent with a molecular subtype in which the ESR1-related gene set was downregulated and the basal type gene set was upregulated, whereas the reverse was true in the gBRCA2 mutants (Supplementary Fig. S6 and Supplementary Table S4). Among the hallmark gene sets, E2F targets, G2M checkpoints, hypoxia, MYC targets, UV response, mitotic spindle, TNFA signaling via NFKB, and glycolysis were upregulated (P < 0.01) in the gBRCA1 mutants, and early estrogen response, bile acid metabolism, pancreas beta cells, and late estrogen response were upregulated in the gBRCA2 mutants (Supplementary Fig. S7).
Gene expression comparison between gBRCA mutation and WT under same IHC subtype

In an extension of the WTS analysis, we conducted comparative analyses to identify the unique gene expression patterns shown with gBRCA mutations under the same IHC subtypes. In our initial analyses, we compared a merged study population with a WT reference group and identified 27 genes with significant differences. Representatively, gBRCA1/2 mutations showed upregulation in CXCL5 and FGFBP1 and downregulation in genes normally expressed in breast tissue, such as TPSAB1, TPSB2, andADIPOQ (Fig. 4D, Supplementary Table S5). Additional functional annotation using Gene Ontology indicated that the 27 DEGs were clustered in serine-type endopeptidase activity and positive regulation of cell proliferation (Supplementary Table S6). Subgroup analyses were conducted by comparing gBRCA1 vs. WT in TNBC patients and gBRCA2 vs. WT in ER-positive patients, and we found that genes and gene sets were expressed differently in each group (Supplementary Fig. S8, Supplementary Tables S7–S10).

Discussion

Since the 1990s, when BRCA1 and BRCA2 were first reported to be associated with hereditary breast cancer (31, 32), their function and clinical utility have been researched extensively. Over the decades, this research has led to the approval of BRCA1/2 mutant–specific clinical treatments (12, 13). With easy access to genomic testing and the development of consensus guidelines for high-risk candidates, an increasing number of patients are being found to have germline BRCA mutations. Consequently, we have questioned the underlying carcinogenesis mechanism triggered by BRCA mutations.
As an initial step, we searched for the underlying co-altered somatic mutation profiles of breast cancer with gBRCA mutations. We emphasize that our results showed a significant co-occurrence of somatic TP53 mutation only with gBRCA1 mutation, not with gBRCA2 mutation. The TP53 pathway plays a pivotal role in maintaining genomic stability through the G1/S checkpoint, which is activated by phosphorylated BRCA1 (33, 34). Because a high incidence rate of somatic TP53 mutation occurs only with gBRCA1 mutation, a second hit alteration in TP53 might trigger the carcinogenesis mechanism in gBRCA1-mutated breast cancer by dysregulating the cell cycles maintained by TP53. In addition, we noticed that the overall somatic PIK3CA mutation status among patients with gBRCA mutations was inconsistent with our knowledge of the general breast cancer population. The mutation rate was significantly lower in breast cancer patients with gBRCA mutations (13.3%) than in those with WT gBRCA (27.7%). This finding is especially evident in ER-positive gBRCA2 patients (11.8%), even though the PIK3CA mutation is associated with hormone receptor-positive breast cancer (35).

Next, we questioned what underlying genomic profile might eventually generate the different molecular subtypes found with gBRCA1 and gBRCA2 mutation (36). As in previous reports, our data demonstrate a high incidence of TNBC with a gBRCA1 mutation and ER-positive breast cancer with a gBRCA2 mutation. As mentioned previously, we thought that the co-alteration of gBRCA1 and TP53, which is enriched in the basal-like subtype (37), might be related to the unique carcinogenesis mechanism of TNBC with a gBRCA1 mutation. In a similar manner, we focused on the genomic instability caused by hormones as a potential mechanism of carcinogenesis with gBRCA2 mutations. In this study, we identified upregulated hormone-related genes, ESR1, AR, and PGR, only with a gBRCA2 mutation. At the same time, it has been reported that the baseline estrogen level is higher than...
normal in those with a gBRCA2 mutation but without cancer (38). Normally, estrogen stimulation binds to the hormone receptor to promote transcription (39), which leads to the frequent formation of RNA-DNA hybrid (R-loop) structures. The abundant formation of R-loops is a replication-stress status that has the potential to cause genomic instability, which is normally prevented by the action of BRCA1 or BRCA2. Therefore, by creating an environment with excess estrogen and the absence of damage control, gBRCA2 deficiency could be a pivotal element in carcinogenesis, which explains the high incidence of hormone receptor–positive breast cancer associated with gBRCA2 mutation. Interestingly, this explains the outcomes from a previous clinical trial, NSABP-P1, in which tamoxifen-treated gBRCA2 carriers showed a 62% reduction in breast cancer incidence compared with the placebo-treated group (40). This supposition should be confirmed through functional validation. As an extension of our analyses, we examined the difference between gBRCA1/2-mutant and WT breast cancer with the same molecular subtype. We found different patterns in gene expression and gene sets between patients with mutant gBRCA and those with the WT, although it was difficult to provide clear clinical implications (Supplementary Fig. S8, Supplementary Table S7–S10).

Finally, we searched for the link between our results and cancer immunotherapy, which has become part of the standard treatment in breast cancer (41). Along with the clinical trials, there has been an effort to develop predictive biomarkers for immunotherapy. Among the candidates, genome-based biomarkers, such as those involved in microsatellite instability (42, 43) and the high TMB caused by DDR pathway alteration, seem to provide an early signal in patient selection for an immune checkpoint inhibitor (ICI) (44, 45). In addition, the potential synergy between blocking DDR pathways, especially using PARP inhibitors, and ICI is under investigation in multiple clinical trials in breast cancer (46). This approach is based on a
previous report showing that PARP inhibition can potentiate DNA damage and lead to inefficient repair, which could lead to the generation of immunologically relevant tumor antigens (47, 48). Unlike our assumption that alteration in the BRCA pathway could increase the total TMB, the median TMB was 6.53 per MB in gBRCA1-mutated patients and 6.55 per MB in gBRCA2-mutated patients, which is no different from the patients with WT BRCA in our dataset and previous reports (49). However, due to our limited number of samples, this finding should also be validated with larger cohorts.

Despite screening a relatively large number of patients, the number of paired samples for the genomic analyses was small and following hypothesis have been presented descriptively. In addition, the method for the genomic analyses were mostly restricted to the somatic alteration. Since this panel is based on tumor sample without matched normal, it was not feasible for the detail analyses such as loss of wild type allele, rearrangement burden, structural alteration and rearrangement pattern as previously evaluated from the Nik-Zainal et al in breast cancer.(50) To elucidate the detail landscape of somatic mutation in germline BRCA mutant in depth analyses using the whole-genome sequencing should be conducted as a future work.

Nonetheless, to our knowledge, this is one of the largest sample sets ever used to comprehensively compare gBRCA mutant samples and the first study to analyze genomic characteristics, including somatic targeted sequencing, for a population of patients with gBRCA mutations, even retrospectively. Moreover, this study includes novel findings, such as discordance between IHC and PAM50 classification in ER-positive patients, which provides the clinical implication that ER-positive IHC patients with a gBRCA mutant subset could have low expression of ESR1, which should be tested using an additional modality.

In conclusion, our results show that breast cancer patients with gBRCA1 and gBRCA2
mutations have distinct genomic backgrounds and gene expression patterns compared with patients with WT gBRCA, and those genomic differences could affect the mechanisms of carcinogenesis in a way that leads to different molecular subtypes.

**Funding**

This work was supported by a grant from the National Research Foundation of Korea (NRF-2018R1A2B6004690) and a grant from the Ministry of Health & Welfare, Republic of Korea (HI13C2096).
References


mutations and their association with hormone receptor expression in breast cancer: a

36. Foulkes WD. Estrogen Receptor Status in BRCA1- and BRCA2-Related Breast Cancer:
The Influence of Age, Grade, and Histological Type 2004;10:2029-34.

37. Silwal-Pandit L, Vollar HKM, Chin SF, Rueda OM, McKinney S, Osako T, et al. TP53
Mutation Spectrum in Breast Cancer Is Subtype Specific and Has Distinct Prognostic

38. Kim J, Oktay K. Baseline E2 levels are higher in BRCA2 mutation carriers: a potential

transcriptional R-loops are the main cause of estrogen-induced DNA damage. Elife
2016;5.

40. King M-C. Tamoxifen and Breast Cancer Incidence Among Women With Inherited
Mutations in BRCA1 and BRCA2. JAMA 2001;286:2251.


Blockade in Tumors with Mismatch-Repair Deficiency. New England Journal of

molecular characterization of clinical responses to PD-1 inhibition in metastatic gastric

al. Nivolumab plus Ipilimumab in Lung Cancer with a High Tumor Mutational Burden.

genes are associated with increased neoantigen burden and a distinct immunophenotype

46. Cesaire M, Thariat J, Candeias SM, Stefan D, Saintigny Y, Chevalier F. Combining
PARP inhibition, radiation, and immunotherapy: A possible strategy to improve the

Association and prognostic significance of BRCA1/2-mutation status with neoantigen
load, number of tumor-infiltrating lymphocytes and expression of PD-1/PD-L1 in high

48. Xiao Y, Freeman GJ. The microsatellite instable subset of colorectal cancer is a
particularly good candidate for checkpoint blockade immunotherapy. Cancer Discov

100,000 human cancer genomes reveals the landscape of tumor mutational burden.
Genome Med 2017;9:34.

of somatic mutations in 560 breast cancer whole-genome sequences. Nature
2016;534:47-54.
Figure legends

Figure 1.
A, Consort flow of study population. B, Sample status of the study population and reference subjects. C, Kaplan-Meier curve of overall survival, recurrence-free survival, and time to distant metastases. D, Nucleotide position of germline BRCA1 and BRCA2 in the study population.

Figure 2.
A, Co-altered somatic mutations in patients with gBRCA1 mutation \( (n = 25) \) and gBRCA2 mutation \( (n = 20) \) with data for additional deep target sequencing (CancerSCAN\textsuperscript{TM}). B, Association between identified somatic mutation and gBRCA1 or gBRCA2 mutations C, and merged gBRCA1/2 mutation. D, Estimated tumor mutation burden using CancerSCAN\textsuperscript{TM}.

Figure 3.
A, Concordance between molecular subtypes identified by immunohistochemistry and PAM50 in patients \( (n = 34) \) with data for whole transcriptome sequencing (WTS). \textit{ESR1} expression is arbitrarily classified as low \( (n = 13) \), intermediate \( (n = 9) \), and high \( (n = 13) \). B, \textit{ESR1} expression profile of patients with WTS data \( (n = 34) \) is visualized based on gBRCA
mutation status as bar graph. In addition, subtypes analyzed by pathologic review and PAM
50 is marked in the bottom of each bar. C, Representative somatic mutation identified in
triple negative subtype in gBRCA1 mutant. The list of genes was either co-occurrent or
mutually exclusive with gBRCA1 which satisfy $P < 0.05$ by fisher’s exact test. D, Similarly,
the list of genes in ER-positive subtype in gBRCA2 mutant showing statistically significant
difference incidence ($P < 0.05$).

Figure 4.

A, Molecular subtype by immunohistochemistry with gBRCA1 ($n = 19$) and gBRCA2 ($n =
15$) mutations with data for whole transcriptome sequencing (WTS). B, Expression profile of
pre-selected representative genes in gBRCA1 and gBRCA2 mutants. C, Differentially
expressed genes (DEGs) between gBRCA1 and gBRCA2. A red dot indicates a gene with
significance. D, DEGs between gBRCA mutants and wild type (WT) patients in merged
population [gBRCA1/2 ($n = 34$) vs. WT ($n = 137$)], in triple negative patients [gBRCA1 ($n =
14$) vs. WT ($n = 65$)], and in ER-positive patients [gBRCA2 ($n = 12$) vs. WT ($n = 78$)]. A red
dot indicates a gene with significance.
Table 1. Baseline demographics of the study population: patients with single BRCA1 and BRCA2 mutations

<table>
<thead>
<tr>
<th>Study population (N = 254)</th>
<th>BRCA1 (n = 128)</th>
<th>BRCA2 (n = 126)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age at diagnosis</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median (range)</td>
<td>40 (23-68)</td>
<td>39 (25-67)</td>
</tr>
<tr>
<td>BRCA mutation*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frameshift</td>
<td></td>
<td></td>
</tr>
<tr>
<td>115 (45.2%)</td>
<td>54 (42.2%)</td>
<td>61 (48.4%)</td>
</tr>
<tr>
<td>Missense</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 (2.0%)</td>
<td>4 (3.1%)</td>
<td>1 (0.8%)</td>
</tr>
<tr>
<td>Nonsense</td>
<td></td>
<td></td>
</tr>
<tr>
<td>110 (43.3%)</td>
<td>53 (41.4%)</td>
<td>57 (45.2%)</td>
</tr>
<tr>
<td>Splicing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18 (7.1%)</td>
<td>12 (9.4%)</td>
<td>6 (4.8%)</td>
</tr>
<tr>
<td><strong>Reason for BRCA test†</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FHx breast or ovary cancer</td>
<td>167 (65.7%)</td>
<td>87 (68.0%)</td>
</tr>
<tr>
<td>FHx other cancer</td>
<td>5 (2.0%)</td>
<td>4 (3.1%)</td>
</tr>
<tr>
<td>Bilateral cancer</td>
<td>54 (21.3%)</td>
<td>27 (21.1%)</td>
</tr>
<tr>
<td>Age under 40</td>
<td>46 (18.0%)</td>
<td>19 (14.8%)</td>
</tr>
<tr>
<td>Breast and ovarian cancer</td>
<td>8 (3.1%)</td>
<td>5 (3.9%)</td>
</tr>
<tr>
<td>Male breast cancer</td>
<td>2 (0.8%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Unidentified</td>
<td>4 (1.6%)</td>
<td>3 (2.3%)</td>
</tr>
<tr>
<td><strong>Menopausal status</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-menopausal</td>
<td>207 (81.2%)</td>
<td>114 (89.1%)</td>
</tr>
<tr>
<td>Post-menopausal</td>
<td>42 (16.5%)</td>
<td>10 (7.8%)</td>
</tr>
<tr>
<td>Other‡</td>
<td>6 (2.4%)</td>
<td>4 (3.1%)</td>
</tr>
<tr>
<td><strong>Prophylactic mastectomy</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prophylactic bilateral oophorectomy</td>
<td>95 (37.4%)</td>
<td>46 (35.9%)</td>
</tr>
<tr>
<td><strong>Surgical record</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MRM or TM</td>
<td>95 (37.4%)</td>
<td>43 (33.6%)</td>
</tr>
<tr>
<td>BCS or PM</td>
<td>159 (62.6%)</td>
<td>85 (66.4%)</td>
</tr>
<tr>
<td><strong>Clinical stage at diagnosis</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>12 (4.7%)</td>
<td>2 (1.6%)</td>
</tr>
<tr>
<td>1A</td>
<td>73 (28.7%)</td>
<td>34 (26.6%)</td>
</tr>
<tr>
<td>1B</td>
<td>4 (1.6%)</td>
<td>1 (0.8%)</td>
</tr>
<tr>
<td>2A</td>
<td>72 (28.4%)</td>
<td>45 (35.2%)</td>
</tr>
<tr>
<td>2B</td>
<td>31 (12.2%)</td>
<td>17 (13.3%)</td>
</tr>
<tr>
<td>3A</td>
<td>37 (14.6%)</td>
<td>20 (15.6%)</td>
</tr>
<tr>
<td>3B</td>
<td>3 (1.2%)</td>
<td>2 (1.6%)</td>
</tr>
<tr>
<td>3C</td>
<td>22 (8.7%)</td>
<td>7 (5.5%)</td>
</tr>
<tr>
<td><strong>Pathology</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Invasive ductal carcinoma</td>
<td>219 (86.2%)</td>
<td>111 (86.7%)</td>
</tr>
<tr>
<td>Other§</td>
<td>35 (13.8%)</td>
<td>17 (13.3%)</td>
</tr>
<tr>
<td>HR (+)</td>
<td>125 (49.2%)</td>
<td>29 (22.7%)</td>
</tr>
<tr>
<td>HR/HER2(+)</td>
<td>10 (3.9%)</td>
<td>4 (3.1%)</td>
</tr>
<tr>
<td>HER2 (+)</td>
<td>3 (1.2%)</td>
<td>3 (2.3%)</td>
</tr>
<tr>
<td>Triple negative</td>
<td>116 (46.7%)</td>
<td>92 (71.9%)</td>
</tr>
</tbody>
</table>

*Some patients are omitted due to limited information about the locus of mutation. †Some patients are counted several times for multiple reasons. ‡Other reasons include previous oophorectomy, total hysterectomy, and unknown menstrual history. §Other types include ductal carcinoma in situ, invasive lobular carcinoma, micropapillary carcinoma, and metaplastic carcinoma.

Abbreviations: FHx, family history; MRM, modified radical mastectomy; TM, total mastectomy; BCS, breast-
conserving surgery; PM, partial mastectomy; HR, hormone receptor
Table 2 Treatment patterns and survival outcomes based on BRCA mutation profile

<table>
<thead>
<tr>
<th>Study population</th>
<th>BRCA1 (n = 128)</th>
<th>BRCA2 (n = 126)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Median duration of follow-up (months)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(N = 254)</td>
<td>55.3</td>
<td>55.7</td>
</tr>
<tr>
<td>(13.7–261.7)</td>
<td>(13.7–237.4)</td>
<td>(14.9–161.7)</td>
</tr>
<tr>
<td><strong>Neoadjuvant chemotherapy</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AC followed by docetaxel</td>
<td>63 (24.8%)</td>
<td>35 (27.3%)</td>
</tr>
<tr>
<td><strong>Adjuvant therapy</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neoadjuvant CTx</td>
<td>197 (77.6%)</td>
<td>92 (71.9%)</td>
</tr>
<tr>
<td>(percent calculated among CTx)</td>
<td>171 (67.3%)</td>
<td>95 (74.2%)</td>
</tr>
<tr>
<td><strong>Adjuvant CTx</strong></td>
<td>197 (77.6%)</td>
<td>92 (71.9%)</td>
</tr>
<tr>
<td>CMF</td>
<td>8 (4.7%)</td>
<td>5 (5.3%)</td>
</tr>
<tr>
<td>FAC</td>
<td>49 (28.7%)</td>
<td>32 (33.7%)</td>
</tr>
<tr>
<td>AC</td>
<td>17 (9.9%)</td>
<td>8 (8.4%)</td>
</tr>
<tr>
<td>AC and taxane</td>
<td>58 (33.9%)</td>
<td>27 (28.4%)</td>
</tr>
<tr>
<td>Other</td>
<td>39 (22.8%)</td>
<td>23 (24.2%)</td>
</tr>
<tr>
<td><strong>Adjuvant hormone therapy</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-op RT</td>
<td>198 (78.0%)</td>
<td>94 (73.4%)</td>
</tr>
<tr>
<td>Recurrence event</td>
<td>132 (52.0%)</td>
<td>33 (25.8%)</td>
</tr>
<tr>
<td>Local recurrence</td>
<td>47 (18.5%)</td>
<td>26 (20.3%)</td>
</tr>
<tr>
<td>Contralateral recurrence</td>
<td>41 (16.1%)</td>
<td>23 (18.0%)</td>
</tr>
<tr>
<td>Systemic recurrence</td>
<td>28 (11.0%)</td>
<td>13 (10.2%)</td>
</tr>
<tr>
<td>Median time to recurrence (months)*</td>
<td>145.0</td>
<td>147.3</td>
</tr>
<tr>
<td>(107.3–170.2)</td>
<td>(127.4–186.2)</td>
<td>(95.2–183.6)</td>
</tr>
<tr>
<td><strong>Palliative chemotherapy</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Survival status</td>
<td>24 (9.5%)</td>
<td>12 (9.4%)</td>
</tr>
<tr>
<td>Deceased</td>
<td>14 (5.5%)</td>
<td>9 (7.0%)</td>
</tr>
</tbody>
</table>

*Median (range)

Abbreviations: AC, adriamycin and cyclophosphamide; CTx, chemotherapy; CMF, cyclophosphamide, methotrexate, and fluorouracil; FAC, fluorouracil, adriamycin, and cyclophosphamide; Op, operation; RT, radiotherapy
Figure 1

A

Total cases of germline BRCA test at SMC (n=2720)

Total samples with positive BRCA test result at SMC (n=386)

Samples with BRCA mutation (n=281)

Study population (N=254)
gBRCA1 (n=128)
gBRCA2 (n=126)

Patients with negative BRCA result (n=2334)

Patient with follow-up period less than 1 year (n=105)

Patients without surgery (n=6)
Limited clinical information (n=16)
Patients with both BRCA1/2 mutation (n=5)

B

gBRCA1 128

25

19

gBRCA2 126

20

15

Reference (Wild Type) 195

C

Overall Survival

Recruent Free Survival

Time to Distant Metastase

D

BRCA1

total mutation frequency : 111 (unique variants : 49)

BRCA2

total mutation frequency : 119 (unique variants : 57)
A. TNBC 16
   HER2+ 1
   ER+ 16
   ER+HER2+ 1

B. Group
   gBRCA1
   gBRCA2

C. ESR1
   SubType/PAM50 (concordance)
   TN /Basal
   HER2+/Her2
   ER+ /LumA
   ER+HER2+/LumB
   − /Normal

D. ARID1B
   BRCA1
   RET
   BRCA2
   PIK3CA

SET
MUTATION TYPE
Frame_Shift_Del
Missense_Mutation
Multi_Hit
In_Frame_Del
Splice_Site
Nonsense_Mutation
Frame_Shift_Ins
Missense_Mutation

IHC SUBTYPE
TN
ER+
Figure 4

A total of 78 genes were significant (red dot)

* Threshold for significance |Exp| > 2 and p-value < 0.01
Clinical characteristics and exploratory genomic analyses of germline BRCA1 or BRCA2 mutations in breast cancer

Sehhoon Park, Eunjin Lee, Seri Park, et al.

Mol Cancer Res  Published OnlineFirst June 17, 2020.

Updated version  
Access the most recent version of this article at: doi:10.1158/1541-7786.MCR-19-1108

Supplementary Material  
Access the most recent supplemental material at: http://mcr.aacrjournals.org/content/suppl/2020/06/17/1541-7786.MCR-19-1108.DC1

Author Manuscript  
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

E-mail alerts  
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions  
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions  
To request permission to re-use all or part of this article, use this link http://mcr.aacrjournals.org/content/early/2020/06/17/1541-7786.MCR-19-1108. Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.