ANTICANCER RESEARCH *40*: 4567-4570 (2020)
doi:10.21873/anticanres.14462

Molecular Profiles of Breast Cancer in a Single Institution

TAKAKO KAMIO[1], HIDENORI KAMIO[1], TAKAKO AOKI[2], YUMIKO ONDO[3], TOMOKI UCHIYAMA[3], KEIKO YAMAMOTO-SHIMOJIMA[2,3], MOTOKO WATANABE[4], TAKAHIRO OKAMOTO[1], HITOSHI KANNO[2] and TOSHIYUKI YAMAMOTO[3]

[1]*Department of Breast and Endocrine Surgery, Tokyo Women's Medical University, Tokyo, Japan;*
[2]*Department of Transfusion Medicine and Cell Processing, Tokyo Women's Medical University, Tokyo, Japan;*
[3]*Institute of Medical Genetics, Tokyo Women's Medical University, Tokyo, Japan;*
[4]*Department of Clinical Genetics, Juntendo University Graduate School of Medicine, Tokyo, Japan*

Abstract. *Background/Aim: Historically, breast cancer has been treated according to an evaluation of biomarkers, such as the estrogen receptor and HER2 status. Recently, molecular profiling has been used to detect driver mutations and select anti-cancer treatment strategies. In addition to detecting pathogenic mutations, the total mutation count (tumor mutation burden) has been considered as another biomarker. Materials and Methods: We performed molecular profiling of 143 breast cancer tissues obtained from resected tissues via surgical operation. Results: Suspected germline mutations were detected in 10% of the patients with a higher somatic mutation ratio. Conclusion: As hypermutated breast cancers are more likely to benefit from certain anti-cancer treatment strategies, molecular profiling can be used as a biomarker.*

All cancers arise as a result of mutations (germ-line/somatic) in the genome (1). Historically, breast cancer has been treated according to a biomarker evaluation, including the estrogen receptor and HER2 status (2, 3). Currently, multigene assays are widely used to predict the risk of relapse after surgery (4). While some of these variabilities are explained by traditional clinicopathological factors (including patient age, tumor stage, histological grade, and estrogen receptor status), molecular profiling studies have defined breast cancer subtypes with distinct clinical outcomes (5).

To better understand the genomic backgrounds of patients, we performed a molecular profiling study of breast cancer patients in a single institution.

Correspondence to: Prof. Toshiyuki Yamamoto, MD, Ph.D., Institute of Medical Genetics, Tokyo Women's Medical University, 8-1 Kawada-cho, Shinjuku-ward, Tokyo 162-8666, Japan. Tel: +81 333538112 ext 24013, e-mail: yamamoto.toshiyuki@twmu.ac.jp

Key Words: Tumor mutation burden, germline mutation, tumor neoantigen.

Materials and Methods

This study was performed in accordance with the Declaration of Helsinki. Additionally, the requisite permissions were obtained from the ethical committee of Tokyo Women's Medical University, and written informed consent was obtained from all participants. The inclusion criteria of this study were that patients were clinically diagnosed with breast cancer and underwent surgical operation between June 2013 and September 2015. The clinical features of the patients were obtained from medical records.

From frozen tissues obtained through surgical operation, the genomic DNA was extracted using the QIAamp DNA extraction kit (QIAGEN, Hilden, Germany), according to the manufacturer's instructions. The purity and concentration of the extracted DNA were measured using a NanoDrop 1000 spectrophotometer (Thermo Scientific, Wilmington, DE, USA) and a Qubit fluorometer (Life Technologies, Carlsbad, CA, USA).

A HaloPlex HS target enrichment system (Agilent Technologies, Santa Clara, CA, USA) was used for this study. The library was designed using SureDesign (6). According to previous studies (1, 7), we selected 59 targeted genes, and all coding exons and intron–exon boundaries of them were included in the panel (Table I). Approximately 57.6 ng of DNA in a volume of 32 μl was used. A260/A280 ratios of 1.8-2.0 were considered acceptable. DNA fragmentation was assessed using agarose gel electrophoresis (2%). The DNA library was constructed using molecular indices and barcodes for simultaneous processing of multiple samples and removal of duplicated reads, respectively. Finally, the enriched library concentration was estimated using the 2200 TapeStation (Agilent Technologies). High-throughput sequencing was performed using 251-bp paired-end reads on a Mi-Seq platform (Illumina, Inc., San Diego, CA, USA), and 12 samples were sequenced at once. The raw data were aligned to the human genome build19. After annotation, FASTQ files were generated automatically. Variant calls were performed using SureCall (Agilent Technologies), and variants were extracted as excel data. Finally, filtering and curation were performed manually. The sample data were omitted from statistical analysis when the percentage of analyzable target regions covered by at least 10 reads was below 90. Prediction scores, including CADD_phred, were obtained using wANNOVAR (8).

Table I. *The list of the genes included in the NGS panel.*

AFF2	*AGTR2*	*AKT1*	*AKT2*	*AL161915.1*	*APC*
ARID1A	*ARID1B*	*ARID2*	*ASXL1*	*ATR*	*BAP1*
BRCA1	*BRCA2*	*CASP8*	*CBFB*	*CCND1*	*CCND3*
CDH1	*CDK4*	*CDKN1B*	*CDKN2D*	*COL12A1*	*DNAH3*
DNAH5	*EGFR*	*ERBB2*	*GATA3*	*HERC2*	*IGF1R*
KMT2C	*KMT2D*	*KRAS*	*LAMA4*	*LDLRAP1*	*MAP2K4*
MAP3K1	*MAP3K13*	*MYC*	*MYH9*	*NCOR1*	*NF1*
PDPK1	*PIK3CA*	*PIK3R1*	*PTEN*	*PTPN22*	*PTPRD*
RB1	*RUNX1*	*SETD2*	*SF3B1*	*SMAD4*	*SMARCD1*
STK11	*STMN2*	*TBX3*	*TP53*	*USH2A*	

Table II. *Suspected germline variants identified in patients.*

	Gene	HGVS (Coding)	HGVS (Protein)	SNP ID	ClinVar evaluation	CADD phred#	Categories	Patient ID
1	*BRCA1*	NM_007294.4:c.3647T>G	p.Leu1216*	rs397509091	Pathogenic		i, ii, iii	27
2	*TP53*	NM_001126112.2:c.1024C>T	p.Arg342*	rs730882029	Pathogenic		i, ii, iii	27, 179
3	*PTEN*	NM_000314.8:c.445C>T	p.Gln149*	rs1060500122	Pathogenic		i, ii	136
4	*NF1*	NM_000267.3:c.3827G>A	p.Arg1276Gln	rs137854556	Pathogenic	36,00	i, iii	169
5	*TP53*	NM_001126112.2:c.329G>C	p.Arg110Pro	rs11540654	Pathogenic	15,79	i, iii	294
6	*BRCA2*	NM_000059.3:c.1705C>T	p.Gln569*				ii	170
7	*CDH1*	NM_004360.5:c.67C>T	p.Gln23*				ii	102
8	*CDH1*	NM_004360.5:c.602_603dup	p.Val202Leufs*14				ii	52
9	*CDH1*	NM_004360.5:c.1585dup	p.Thr529Asnfs*8				ii	176
10	*PTEN*	NM_000314.8:c.569delC	p.Pro190Glnfs*9				ii	28
11	*PTEN*	NM_000314.8:c.1007_1011delACTTT	p.Tyr336Phefs*5				ii	139
12	*TP53*	NM_001276760.2:c.889G>T	p.Glu297*				ii	116
13	*BRCA2*	NM_000059.4:c.9190G>C	p.Asp3064His			17,46	iii	15
14	*BRCA1*	NM_007294.4:c.4096G>T	p.Gly1366Cys			21,00	iii	133

HGVS: Human Genome Variation Society nome clature; #Only variants with missense alterations show CADD_phred.

Germline variants were assumed to be observed in the NCCN guideline germline list (*BRCA1, BRCA2, CDH1, NF1, PTEN, STK11*, and *TP53*) (9). The filtering policy used was that each variant had to fulfill the following three conditions: a) a variant frequency below 1% in ExAC_EAS (The Genome Aggregation Database [former name Exome Aggregation Consortium (ExAC)]) (10), 1000G_EAS and ALL (1000 Genomes) (11), HGVD (Human Genetic Variation Database) (12), and dbSNP (13); b) a higher than 30% variant allele frequency in the total depth; c) should meet at least one of the following categories: i) variants registered in ClinVar (14) with "pathogenic" or "likely pathogenic", ii) variants related to loss of function (nonsense or frameshift), or iii) *BRCA1/2* variants not registered in ClinVar and HGMD, but a CADD_phred was higher than 15.

Somatic mutations were defined as follows: a) variant frequencies below 1% in ExAC_EAS, 1000G_EAS, 1000G_ALL, HGVD, and dbSNP; b) variants including any nonsynonymous, frameshift, inframe, and splicing acceptor/donor site mutations; c) variant allele frequency in the total depth of 25%-50%; d) a variant read depth of higher than 10; and e) CADD_phred higher than 15.

We analyzed the relationship between the clinical features of patients and the results of molecular analyses using the Wilcoxon signed-rank test and the Jonckheere–Terpstra trend test. Statistical analyses were performed using R (15). A *p*-value lower than 0.05 was considered statistically significant.

Results

Patients. The 142 tissues from 142 patients (their average ages at this study were 63.1±13.4; range=30-93 years) were included in this study.

Molecular analyses. In two samples, the percentage of analyzable target regions covered by at least 10 reads was below 90; thus, these two samples were removed from the statistical analysis. The minimum and maximum numbers of the total variant call in all samples were 112 and 157, respectively. After filtering, 303 variant calls (193 variant types in 43 genes) remained (Data are available upon request). One of the variants in *TP53* was drug treatable/actionable. Among the identified variants, 14 types of variants were considered as germline variants, owing to the higher allele frequencies in the read depth (Table II). One patient had two suspected germline variants.

Statistical analysis. Fourteen patients with suspected germline variants had 4.5 variants on average. In contrast, 126 patients showed 1.9 variants on average. There was a

Table III. *Comparison based on suspected germline mutations.*

	Patients without suspected germline mutations	Patients with suspected germline mutations	Results
Numbers of patients	126	14	Wilcoxon test
Average number of somatic variants	1.9	4.5	$p<0.0001$
Standard deviation	1.4	1.5	
Minimum	0	2	*t*-test
Median	2	4.5	$p<0.0001$
Maximum	6	7	

Table IV. *Results of statistical analyses of total variant call numbers.*

Total variant calls	0	1	2	3	≥4	*p*-Value
N	27	31	27	24	31	
Age at diagnosis (years old)						
N	14	17	18	18	15	0.031
Average	58.6	60.4	66.2	65.2	64.0	
S.D.	16.1	14.3	11.7	11.4	14.5	
Minimum	30	44	42	45	43	
Median	61	55	68.5	69	66	
Maximum	84	84	93	81	84	
Unknown	13	14	9	6	16	
Menopause						
Before	4	7	9	4	7	0.6148
After	17	16	15	18	17	
During	1	1	0	0	2	
Unknown	5	7	3	2	5	
Stages (post-operation UICC)						
I	9	14	10	9	8	0.2478
IIA	6	5	4	7	8	
IIB	2	2	3	3	5	
IIIA	3	3	5	1	3	
IIIB	1	0	1	1	0	
IIIC	0	0	0	1	0	
IV	0	0	1	0	0	
Unknown	6	7	3	2	7	
Number of positive lymph node metastases						
0	19	21	13	15	20	0.3281
1	2	0	2	3	4	
2	2	3	4	0	0	
3	0	0	2	1	1	
4	1	0	1	1	1	
≥5	1	3	2	2	2	
Unknown	2	4	3	2	3	
BMI						
N	24	27	23	23	29	0.5893
Average	22.5	22.4	22.7	22.8	21.9	
S.D.	3.1	4.1	5.1	3.7	3.2	
Minimum	15.0	17.5	15.2	16.1	16.8	
Median	22.2	21.8	22.7	23.0	21.5	
Maximum	27.4	35.9	37.4	33.1	29.0	
Unknown	3	4	4	1	2	
Maximum diameter of cancer suspected from radiological images (cm)						
N	25	27	25	21	28	0.7624
Average	3.8	2.5	6.5	6.0	2.8	
S.D.	2.3	1.3	11.3	15.4	1.2	
Minimum	1.3	1.0	1.1	1.0	0.9	
Median	3.2	2.1	2.8	2.3	2.8	
Maximum	10.4	7.0	53.1	72.8	5.0	
Unknown	2	4	2	3	3	

S.D.: Standard deviation; BMI: body mass index.

statistically significant difference between the two groups (Table III). Regarding the clinical features, there was no statistical significance in correlation with the number of total variants (Table IV).

Discussion

In this study, 14 types of suspected germline variants were identified in 14 samples from 140 patients, indicating that 10% of the patients showed possible hereditary/familial cancers. Four patients showed suspected *BRCA1/2* germline variants. It is known that mutations in *BRCA1/2* are responsible for the hereditary breast and ovarian cancer (HBOC) syndrome (16). *TP53* variants were identified in three patients. This gene is related to Li-Fraumeni syndrome. Breast cancer can occur in patients with Li-Fraumeni syndrome. Similarly, *PTEN, NF1*, and *CDH1* can also be related to familial breast cancer (17). Patients with suspected germline variants showed higher statistically significant variant calls. This is reasonable because loss of function of tumor suppressor genes causes multiple gene mutations.

Recently, it has been considered that the more somatic mutations a tumor has, the more neoantigens it is likely to form (18). This condition is known as tumor mutational burden (TMB) and can represent a useful estimation of the tumor neoantigenic load. TMB is an emerging biomarker in cancer and has been associated with microsatellite instability, defective DNA replication/repair, and response to programmed cell death -1 (PD-1) and programmed death-ligand 1 (PD-L1) blockade immunotherapy (19). As hypermutated breast cancers are more likely to benefit from PD-1 inhibitors, mutation analysis can be used as a biomarker (20).

There are some limitations in this study. Because this study was not designed to analyze germline DNA extracted from normal tissues, including blood, it was impossible to confirm true germline mutations. The study design was planned in 2016. At that time, driver genes for breast cancer were not fully understood and some important genes were not included in the gene panel. Therefore, most of the conclusions were drawn from the information available by our study design. However, the final conclusion does not contradict previous studies.

Conflicts of Interest

The Authors declare they have no conflicts of interest regarding this study.

Authors' Contributions

TK, HK and TY designed the study. TA, YO, and BW carried out all experiments. TY supervised all experiments performed. TU, KY-S and MW participated in the data interpretation. TK drafted the manuscript. TY prepared the final manuscript. As the principal investigator, TO supervised the study. All Authors read and approved the final manuscript.

References

1 Stephens PJ, Tarpey PS, Davies H, Van Loo P, Greenman C, Wedge DC, Nik-Zainal S, Martin S, Varela I, Bignell GR, Yates LR, Papaemmanuil E, Beare D, Butler A, Cheverton A, Gamble J, Hinton J, Jia M, Jayakumar A, Jones D, Latimer C, Lau KW, McLaren S, McBride DJ, Menzies A, Mudie L, Raine K, Rad R, Chapman MS, Teague J, Easton D, Langerod A, Lee MT, Shen CY, Tee BT, Huimin BW, Broeks A, Vargas AC, Turashvili G, Martens J, Fatima A, Miron P, Chin SF, Thomas G, Boyault S, Mariani O, Lakhani SR, van de Vijver M, van't Veer L, Foekens J, Desmedt C, Sotiriou C, Tutt A, Caldas C, Reis-Filho JS, Aparicio SA, Salomon AV, Borresen-Dale AL, Richardson AL, Campbell PJ, Futreal PA and Stratton MR: The landscape of cancer genes and mutational processes in breast cancer. Nature *486*: 400-404, 2012. PMID: 22722201. DOI: 10.1038/nature11017

2 Papadimitriou M, Mountzios G and Papadimitriou CA: The role of parp inhibition in triple-negative breast cancer: Unraveling the wide spectrum of synthetic lethality. Cancer Treat Rev *67*: 34-44, 2018. PMID: 29753961. DOI: 10.1016/j.ctrv.2018.04.010

3 Petrelli F, Tomasello G, Barni S, Lonati V, Passalacqua R and Ghidini M: Clinical and pathological characterization of HER2 mutations in human breast cancer: A systematic review of the literature. Breast Cancer Res Treat *166*: 339-349, 2017. PMID: 28762010. DOI: 10.1007/s10549-017-4419-x

4 Naito Y and Urasaki T: Precision medicine in breast cancer. Chin Clin Oncol *7*: 29, 2018. PMID: 30056731. DOI: 10.21037/cco.2018.06.04

5 Baird RD and Caldas C: Genetic heterogeneity in breast cancer: The road to personalized medicine? BMC Med *11*: 151, 2013. PMID: 23800221. DOI: 10.1186/1741-7015-11-151

6 Agilent SureDesign by Agilent Technologies. Available at: https://earray.chem.agilent.com/suredesign/?locale=ja [Last accessed June 24, 2020]

7 Nik-Zainal S, Davies H, Staaf J, Ramakrishna M, Glodzik D, Zou X, Martincorena I, Alexandrov LB, Martin S, Wedge DC, Van Loo P, Ju YS, Smid M, Brinkman AB, Morganella S, Aure MR, Lingjaerde OC, Langerod A, Ringner M, Ahn SM, Boyault S, Brock JE, Broeks A, Butler A, Desmedt C, Dirix L, Dronov S, Fatima A, Foekens JA, Gerstung M, Hooijer GK, Jang SJ, Jones DR, Kim HY, King TA, Krishnamurthy S, Lee HJ, Lee JY, Li Y, McLaren S, Menzies A, Mustonen V, O'Meara S, Pauporte I, Pivot X, Purdie CA, Raine K, Ramakrishnan K, Rodriguez-Gonzalez FG, Romieu G, Sieuwerts AM, Simpson PT, Shepherd R, Stebbings L, Stefansson OA, Teague J, Tommasi S, Treilleux I, Van den Eynden GG, Vermeulen P, Vincent-Salomon A, Yates L, Caldas C, van't Veer L, Tutt A, Knappskog S, Tan BK, Jonkers J, Borg A, Ueno NT, Sotiriou C, Viari A, Futreal PA, Campbell PJ, Span PN, Van Laere S, Lakhani SR, Eyfjord JE, Thompson AM, Birney E, Stunnenberg HG, van de Vijver MJ, Martens JW, Borresen-Dale AL, Richardson AL, Kong G, Thomas G and Stratton MR: Landscape of somatic mutations in 560 breast cancer whole-genome sequences. Nature *534*: 47-54, 2016. PMID: 27135926. DOI: 10.1038/nature17676

8 wANNOVAR. Available at http://wannovar.wglab.org/ [Last accessed on June 24, 2020]

9 Daly MB, Pilarski R, Berry M, Buys SS, Farmer M, Friedman S, Garber JE, Kauff ND, Khan S, Klein C, Kohlmann W, Kurian A, Litton JK, Madlensky L, Merajver SD, Offit K, Pal T, Reiser G, Shannon KM, Swisher E, Vinayak S, Voian NC, Weitzel JN, Wick MJ, Wiesner GL, Dwyer M and Darlow S: Nccn guidelines insights: Genetic/familial high-risk assessment: Breast and ovarian, version 2.2017. J Natl Compr Canc Netw *15*: 9-20, 2017. PMID: 28040716. DOI: 10.6004/jnccn.2017.0003

10 gnomAD: Available at: https://gnomad.broadinstitute.org [Last accessed June 24, 2020]

11 1000 Genomes. Available at https://www.internationalgenome.org [Last accessed June 24, 2020]

12 Human Genetic Variation Database. Available at: http://www.hgvd.genome.med.kyoto-u.ac.jp/ [Last accessed June 24, 2020]

13 dbSNP. Available at: https://www.ncbi.nlm.nih.gov/snp [Last accessed June 24, 2020]

14 ClinVar. Available at: https://www.ncbi.nlm.nih.gov/clinvar [Last accessed June 24, 2020]

15 The Comprehensive R Archive Network. Available at: https://cran.r-project.org/ [Last accessed June 24, 2020]

16 Couch FJ, Shimelis H, Hu C, Hart SN, Polley EC, Na J, Hallberg E, Moore R, Thomas A, Lilyquist J, Feng B, McFarland R, Pesaran T, Huether R, LaDuca H, Chao EC, Goldgar DE and Dolinsky JS: Associations between cancer predisposition testing panel genes and breast cancer. JAMA Oncol *3*: 1190-1196, 2017. PMID: 28418444. DOI: 10.1001/jamaoncol.2017.0424

17 Sheikh A, Hussain SA, Ghori Q, Naeem N, Fazil A, Giri S, Sathian B, Mainali P and Al Tamimi DM: The spectrum of genetic mutations in breast cancer. Asian Pac J Cancer Prev *16*: 2177-2185, 2015. PMID: 25824734. DOI: 10.7314/apjcp.2015.16.6.2177

18 Chan TA, Yarchoan M, Jaffee E, Swanton C, Quezada SA, Stenzinger A and Peters S: Development of tumor mutation burden as an immunotherapy biomarker: Utility for the oncology clinic. Ann Oncol *30*: 44-56, 2019. PMID: 30395155. DOI: 10.1093/annonc/mdy495

19 Hatakeyama K, Nagashima T, Urakami K, Ohshima K, Serizawa M, Ohnami S, Shimoda Y, Ohnami S, Maruyama K, Naruoka A, Akiyama Y, Kusuhara M, Mochizuki T and Yamaguchi K: Tumor mutational burden analysis of 2,000 japanese cancer genomes using whole exome and targeted gene panel sequencing. Biomed Res *39*: 159-167, 2018. PMID: 29899191. DOI: 10.2220/biomedres.39.159

20 Barroso-Sousa R, Jain E, Cohen O, Kim D, Buendia-Buendia J, Winer E, Lin N, Tolaney SM and Wagle N: Prevalence and mutational determinants of high tumor mutation burden in breast cancer. Ann Oncol *31*: 387-394, 2020. PMID: 32067680. DOI: 10.1016/j.annonc.2019.11.010

Received June 5, 2020
Revised June 26, 2020
Accepted June 29, 2020