Original Study

Identifying Genomic Alterations in Stage IV Breast Cancer Patients using MammaSeq™: An International Collaborative Study

Osama Shiraz Shah^{1,2}, Atilla Soran³, Mustafa Sahin⁴, Serdar Ugras⁴, Esin Celik⁴, Peter C. Lucas^{5,6}, Adrian V. Lee^{1,6,7,8}

1 Women's Cancer Research Center, UPMC Hillman Cancer Center, Pittsburgh, PA, USA

2 Integrative Systems Biology Program, University of Pittsburgh, Pittsburgh, PA, USA

3 UPMC Department of Surgery, Pittsburgh, PA, USA

4 School of Medicine, Selçuk University, Konya, Turkey

5 Department of Pathology, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA

6 UPMC Hillman Cancer Center, Pittsburgh, PA, USA

7 Department of Pharmacology & Chemical Biology, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA

8 Magee-Womens Research Institute, Pittsburgh, PA, USA

Correspondence to:

Adrian V. Lee, Ph.D.
UPMC Hillman Cancer Center
204 Craft Avenue
Pittsburgh
PA 15213
T: 412-641-7557
F: 412-641-2458

Conflict of interest statement

The authors indicated no potential conflicts of interest

Highlights

- 41 stage IV stage breast cancer patients of Turkish descent were sequenced using MammaSeq™
- 49 single nucleotide variations and 10 copy number variations identified
- *PIK3CA* and *TP53* mutations were present in 24% and 17% of the samples respectively
- 37% of the samples had *ERBB2/GRB7* gains and 7% had loss of *BRCA2/RB1* locus
- Eight clinically significant alterations were identified

Micro Abstract

We performed targeted sequencing using DNA from FFPE samples of 41 stage IV breast cancer patients using MammaSeqTM, a breast cancer gene specific targeted sequencing panel. In total, 49 single nucleotide variations (SNVs) and 10 copy number variations (CNVs) were identified. Eight alterations (3 CNVs – *ERBB2*, *FGFR1* and *AR* copy number gains and 5 SNVs – *IDH1*.R132H, *TP53*.E204*, *PI3KCA*.E545K, *PI3KCA*.H1047R and *PI3KCA*.R88Q) were identified to have clinical significance by PMKB and OncoKB databases.

ABSTRACT

Background

Identification of genomic alterations present in cancer patients may aid in cancer diagnosis and prognosis and may identify therapeutic targets. In this study, we aimed to identify clinically actionable variants present in stage IV breast cancer (BC) samples.

Materials and Methods

DNA was extracted from formalin fixed paraffin embedded (FFPE) samples of BC (n=41). DNA was sequenced using MammaSeqTM, a BC specific next generation sequencing panel targeting 79 genes and 1369 mutations. Ion Torrent Suite 4.0 was used to make variant calls on the raw data and the resulting single nucleotide variants were annotated using CRAVAT toolkit. SNVs were filtered to remove common polymorphisms and somatic variants. CNVkit was employed to identify copy number variations. The Precision Medicine Knowledgebase (PMKB) and OncoKB Precision Oncology Database were used to associate clinical significance with the identified variants.

Results

A total of 41 Turkish BC patient samples were sequenced (read depth of 94 – 13340, median of 1529). These samples were from patients diagnosed with various BC subtypes including invasive ductal carcinoma (IDC), invasive lobular carcinoma (ILC), apocrine BC and micropapillary BC. In total, 59 different alterations (49 SNVs and 10 CNVs) were identified. From these, 8 alterations (3 CNVs – *ERBB2*, *FGFR1* and *AR* copy number gains and 5 SNVs – *IDH1*.R132H, *TP53*.E204*, *PI3KCA*.E545K, *PI3KCA*.H1047R and *PI3KCA*.R88Q) were identified to have some clinical significance by PMKB and OncoKB. Moreover, the top five genes with most SNVs included *PIK3CA*, *TP53*, *MAP3K1*, *ATM* and *NCOR1*. Additionally, copy number gains and losses were found in *ERBB2*, *GRB7*, *IGFR1*, *AR*, *FGFR1*, *MYC* and *IKBKB*, and *BRCA2*, *RUNX1* and *RBI* respectively.

Conclusion:

We identified 59 unique alterations in 38 genes in 41 stage IV BC tissue samples using MammaSeqTM. Ten of these alterations were found to have some clinical significance by OncoKB and PKMB. This study highlights the potential use of cancer specific NGS panels in clinic to get better insight into the patient-specific genomic alterations.

Keywords

Advance breast cancer, targeted next generation sequencing, NGS panel, MammaSeqTM

INTRODUCTION

Next generation sequencing (NGS), with its improving speed, accuracy and cost, is becoming increasingly feasible in clinical care. However, current selection of therapies for breast cancer (BC) patients are primarily based on clinical and histological factors and genomic factors such as genomic heterogeneity are often overlooked¹. Supplementing therapy selection with NGS diagnostics can help account for genomic heterogeneity and aid in designing effective treatment regimens for BC patients. NGS based diagnostics can be especially useful for management of stage IV BC disease which have a very heterogenous genetic landscape². For such cases, NGS can identify which genomic features are/could be responsible for drug resistance and what alternate clinically actionable targets might be available.

In this study we investigated the genomic alteration landscape of stage IV BC patients of Turkish descent (Table 1) by sequencing formalin-fixed paraffin-embedded (FFPE) BC samples using MammaSeqTM³. We identified 59 alterations of which 8 were clinically actionable. With this study we provide additional support for use of NGS diagnostics in management of stage IV BC patients.

MATERIALS AND METHODS

Patients and Samples Processing

Following review and approval by the Institutional Review Board at the University of Pittsburgh, we obtained FFPE tissues from BC cases diagnosed at Selcuk University, Konya, Turkey ([Table 1, supplementary File 1](#)). Samples were sectioned and reviewed by a breast pathologist at UPMC Magee Women's Hospital (PCL). Sections were macro-dissected followed by deparaffinization (Qiagen cat#19093) and DNA isolation (QIAamp DNA FFPE kit, cat#56404), as per manufacturer's instructions. Isolated DNA was quantified using the ThermoFisher Qubit dsDNA S/BR kit.

Ion Torrent Sequencing

Samples with adequate DNA were sent for Ion Torrent Sequencing at the University of Pittsburgh Genomics Core following established methods³. In summary, we collected 20ng of DNA (divided equally between two pools of primers) for library preparation using Ion AmpliSeq™ library kit (ThermoFisher Scientific) and MammaSeq™ primer panel. Emulsion PCR was used for target enrichment using Ion OneTouch 2 system (ThermoFisher Scientific). Enriched library was loaded onto Ion chips and sequenced on Ion Torrent Personal Genome Machine (PGM™, ThermoFisher Scientific).

Quality control analysis

TarSeqQC⁴, a R package for quality control analysis of targeted NGS experiments, was used for quality control analysis of the generated NGS dataset. Coverage of amplicons sequenced by primer mix 1 (pool 1) and primer mix 2 (pool2) were assessed.

Variant calling

Variant calls were performed using Ion Torrent Suite 4.0⁵. Briefly, the raw fastq files were aligned to hg19 reference genome to generate VCF 4.0 files. The variants were prefiltered for allele frequency cutoff of 5%. For further variant filtering we followed established methods as described in Smith et al 2019³. Briefly, we annotated the variants using Cravat CHASM v4.3⁶ removed variations which were: 1) synonymous, 2) present in 30% or more of the samples, 3) having high strand bias, 4) found in gnomAD or 1000 genomes database with an allele frequency of more than 1%, 5) having allele frequency of more than 90% and 6) having a read depth of 50 or lower.

To compute CNVs, we used the Copy Number Ratio (CNR) outputs from CNVkit⁷. CNR files contain following: $\log_2 \frac{\text{Amplicon reads in tumor}}{\text{Amplicon reads in pooled normals}}$ value for all gene amplicons. CNR value residing in each bin were aggregated and used to calculate a copy number value for each gene in that bin using the formula: $2 * 2^{(\text{aggregated CNR})}$. CNR values with read depth lower than 5 and those that lied 2 standard deviations away from mean in each sample to discard low confidence amplicons and potential outliers. For detailed information on these calculations see CNVkit⁷ documentation (<https://cnvkit.readthedocs.io/en/stable/pipeline.html>).

Clinical Actionability Analysis

We used PMKB⁸ and OncoKB to associate clinical significance to the shortlisted CNVs and SNVs. Briefly, we downloaded the databases and then cross checked the presence of each alteration in it.

Data and code:

R⁹ was used for quality control, variant annotation, filtering and clinical significance association as well as for making the barplots, venn diagram (VennDiagram¹⁰) and oncoprint (Complex Heatmaps¹¹). The code is available openly at on github:

https://github.com/osamashiraz/leeoesterreich_MammaSeq-Turkish.Cohort/.

RESULTS AND DISCUSSION

Cohort Overview and Sequencing

For this study, we obtained FFPE tissue blocks from BC cases diagnosed at Selcuk University, Konya, Turkey. The mean age of patients was 50.5 with standard deviation of 13.1 (Table 1). Tumor purity (“Purity”) values are based on the interpretations of the pathologist. ER, PR and HER2 status listed in Table 1 are based on Immuno-HistoChemistry (IHC) test. Most samples were ER+, PR- and HER2+. 41.5% of the samples were from bone metastasis tissue, 22% from mixed/complex tissue, 12.2% from lung and lymph node sites each, 5% from brain and pleura each and 2.4% from liver. Most of the samples (87.8%) were diagnosed as invasive ductal carcinoma (IDC), 7% were invasive lobular carcinoma (ILC) and few (5%) could not be diagnosed and were characterized as OTHER/UNKOWN (see Supplementary File 1 for details).

Alteration Landscape

In total, all 41 samples were successfully sequenced using MammaSeq™ panel with mean coverage ranging from 94-133,340, average mean coverage of 2,308 and median of 1,529 (Supplementary Fig. 1). Variant calls were made using IonTorrent 4.0⁵ and copy number calls using CNVkit⁷. SNVs from variant calls were annotated using Cravat CHASM v4.3⁶ and were filtered to remove common polymorphisms and sequencing artifacts (see Supplementary Fig. 3) as described in Materials and Methods section. CNV computation was performed using CNR values from CNVkit⁷ as described in Materials and Methods. Gains were defined as copy number > 6 and losses as copy number < 1. During quality control, we noticed that amplicons sequenced by primer mix 2 (pool 2) had a lower overall coverage compared to primer mix 1 (pool 1) (Supplementary Fig. 2). To account for the influence of low coverage amplicons on variant identification and copy number computation, we removed low read depth variant calls from SNV analysis and low coverage amplicons from CNV analysis as mentioned in Materials and Methods section (also see Supplementary File 2). The final oncoprint showing the genomic alterations detected by MammaSeq™ is shown in Fig. 1.

We found a total of 59 unique alterations. Forty-nine of these were SNVs. *PIK3CA* gene is an important regulator of PI3K-AKT pathway which controls cell proliferation and survival¹². In our analysis *PIK3CA* was the most frequently mutated gene. *PIK3CA* mutations were seen at rate of 27% in our cohort while other metastatic BC studies report 28%¹³, ~29%¹⁴ and ~30% (MBC project - The Metastatic Breast Cancer Project (<https://www.mbcproject.org/>), a project of Count Me In (<https://joincountmein.org/>)). *PIK3CA* mutation rates seen in stage IV/metastatic BCs are lower than those reported by primary BC studies (~35% in TCGA¹⁵ and 37% in METABRIC^{16,17}). Regardless of these differences, *PIK3CA* mutations are widespread in primary and metastatic BCs. The FDA recently approved the first *PIK3CA* inhibitor, Alpelisib, for patients with *PIK3CA* mutation confirmed via a companion diagnostic¹⁸.

TP53 is an important tumor suppressor which inhibits abnormal proliferation¹⁹. This was the second most frequently mutated gene in our cohort. The SNV rate in *TP53* was 17% while other metastatic BC studies have reported 13%¹³, ~38%¹⁴ and 30% (MBC project). Both our and Pezo et al’s¹³ study use a targeted NGS approach; i.e. only hotspot regions of select genes were sequenced to identify SNVs. Targeted approaches are not comprehensive and could underrepresent the net rate of mutations, especially for tumor suppressor genes such as *TP53* where pathogenic mutations may be spread throughout the coding region. In contrast, the Lefebvre et al¹⁴ and MBC project use whole exome sequencing which offers more comprehensive SNV identification by covering all gene exons. This might explain why the latter two reports show higher numbers of *TP53* mutations. Furthermore, *TP53* gene mutation rate in primary BC studies was ~34%^{15–17} which overlaps with the rates seen in metastatic BCs.

MAPK31 is MAPK kinase protein that activates the MAPK/ERK pathway to control proliferation, migration and other cell fates²⁰. We found *MAPK31* to be mutated at a rate of 15% compared

to ~8%¹⁴ and ~4% (MBC project) in other metastatic BC studies. While in primary BC studies the mutation rate was ~8%¹⁵ and ~9%^{14,16}.

ATM is a tumor suppressor which is important for DNA repair²¹. We observed the *ATM* gene to be mutated at a rate of 10%. Others have reported ~4%¹⁴ and ~2% (MBC project) for metastatic BCs. Apart from these frequently mutated genes we also found less frequent mutations in *AR*, *CCNE1*, *FGFR1*, *NCOR1*, *PTCH1*, *ARID1A*, *AURKB*, *CCND1*, *CDKN1B*, *CDKN2B*, *CTCF*, *EGFR*, *ERBB4*, *FGF19*, *FGFR2*, *FGFR4*, *FOXA1*, *GATA3*, *IDH1*, *IKBKE*, *INPP4B*, *JAK3*, *MAP2K4* and *MET* (see Fig. 1 and supplementary file 3). Oncoprint with frequency of each unique alteration is shown in Supplementary Fig. 4.

CNVs are an important class of genomic alterations. We identified 10 unique CNVs. Most common of these was *ERBB2/GRB7* locus gains. *ERBB2* is a receptor tyrosine kinase and belongs to the EGFR family. Aberrant activation of *ERBB2* is associated with more aggressive BCs and reduced patient survival²². We found *ERBB2/GRB7* gain in 37% of the cases compared to 22.36% in MBC project cohort. These values are higher than those reported for primary BCs¹⁵⁻¹⁷ (~12-13%). Moreover, in our previous study we observed 2-fold higher expression of *ERBB2* gene in 35% of the brain metastases when compared to matched primary BC tissue²³. These observations suggest that *ERBB2* gains may be frequently acquired as tumor progresses to metastatic disease. This makes it critical to examine the HER2 status of stage IV/metastatic patients to determine if HER2 therapies could be of therapeutic value.

We identified deletions of *BRCA2/RB1* locus at 7% rate (2/3 of these cases had ILC subtype). It is important to note that loss of *BRCA2* is associated with resistance to DNA damaging agents²⁴. Similarly, *RB1* deficient tumors could show resistance to CDK4 inhibitors¹⁴. Hence, *RB1/BRCA2* locus genotyping in stage IV/metastatic BC patients can be of prognostic importance as it may help determine if cancer therapies such as CDK4/6 inhibitors would be effective or not.

RUNX1 is an important cell fate regulator of ER+ mammary luminal cells and is often mutated in BCs^{15,25}. *RUNX1* gains were seen in 7% of the samples. *RUNX1* amplifications are rare in both primary BCs¹⁵⁻¹⁷ (~1-2%) and metastatic BCs¹⁴ (~3%). Moreover, its role in BC tumorigenesis or metastasis has not been very well established. However, *RUNX1* upregulation in colorectal cancer has been reported to drive metastasis via WNT pathway²⁶. Further studies will be needed to validate our findings on *RUNX1* amplifications and their role in metastases.

IGF1R controls the insulin-like growth factor pathway, an important regulator of survival, motility and proliferation²⁷. We also found *IGF1R* gains in 5% of the samples. This frequency was similar to that reported for primary BCs i.e. ~3-5%¹⁵⁻¹⁷ and metastatic BCs ~4%¹⁴. Based on these findings, *IGF1R* gains seem to have similar frequency in primary and metastatic BCs.

We also observed gains in *IKBKB*, *MYC*, *AR* and *FGFR1* at ~2% frequency each. Except for *AR* gains, rest of the gene gains were reported at a much lower frequency by Lefebvre et al in metastatic BCs¹⁴ (*IKBKB*: ~6%, *MYC*: ~19%, and *FGFR1*: ~13%). There could be few reasons for these differences. 1) As MammaSeqTM is a targeted panel and it does not cover the entire genes it targets which prevents comprehensive CNV discovery. 2) Also, the cohort is of small size and it does not capture the complete heterogeneity of stage IV/metastatic BC cases.

Clinical Actionability of Identified Variant

Next, we sought to investigate the clinical actionability of alterations detected by MammaSeqTM. For this purpose, we used two precision oncology databases; OncoKB²⁸ and PMKB⁸. Eight alterations were highlighted by these databases (4 were reported only by PMKB, 1 only by OncoKB and 3 by both) (See Fig. 2). Variants of high clinical significance included *PIK3CA*-E545K mutation, *ERBB2* gain and *AR* gain. Among these, only *ERBB2* gain was commonly detected by both databases, while the other two were detected only by PMKB (Table 2). Other clinically actionable variants included: *FGFR1* gain (associated

with endocrine resistance²⁹), *IDH1*-R132H (associated with increased cell migration in gliomas³⁰), *TP53*-E204*, *PIK3CA*-H1047R and *PIK3CA*-R88Q. All the identified *PI3K* mutations are activating in nature³¹. Among these, the most common alterations were *ERBB2* gain (37%), *PIK3CA*-H1047R (20%), and *PIK3CA*-E545K (7%) (see [Supplementary Fig. 4](#)). While remaining 5 alterations had a frequency of ~2%.

FDA recently approved the first *PIK3CA* inhibitor, Alpelisib, for treating *PI3KCA* mutant BCs¹⁸. *ERBB2* gains are seen in both primary BCs and metastatic BCs with a somewhat higher rate in the latter. Antibody-based, *ERBB2*-targeting therapeutics, including trastuzumab, have proven to be effective for these patients³². Thus, it is important to screen metastatic BC patients for *ERBB2* gains and provide treatment where applicable. Previously, it has been suggested that high *AR* expression has been associated with better response to endocrine-therapy in ER+ BCs³³. However, recent data from BIG-1-98 did not report significant association of *AR* expression with prognosis or endocrine treatment differences³⁴. Finally, FGFR family inhibitors for BCs are in testing and show promise³⁵.

CONCLUSION

With this study, we highlight the potential use of cancer specific NGS panels in diagnostic management of cancer patients in clinic. Such panels can be used in clinic to provide insight into the genomic alteration landscape of cancer patients and guide patient specific cancer therapy.³⁶

CLINICAL PRACTICE POINTS

Therapy selection for BC patients is primarily based on clinical and histological factors. However, genomic features such as genomic heterogeneity of BC play an important role in predicting treatment response. NGS diagnostic assays are being widely developed and tested if they can improve conventional diagnostic and prognostic assays for different cancers. Such panels have been tested in BC. However, they have been limited to primary BC disease. We recently developed MammaSeqTM as an NGS based diagnostic/prognostic assay for both primary and metastatic BCs. In this study, we use MammaSeqTM to decode the underlying genomic heterogeneity of metastatic BC patients of Turkish descent and identify clinically actionable alterations.

We identified 59 unique alterations (49 SNVs and 10 CNVs) in 38 genes. *ERBB2* gains were the most frequent alteration and are targetable using anti-*HER2* drugs. Given the high frequency of *ERBB2* gains in metastatic BCs compared to primary BCs, it is critical to assess *HER2/ERBB2* status in metastatic BCs and treat accordingly. Moreover, various activating PI3K mutations were found. Patients harboring such mutations can benefit from Alpelisib (a recent FDA approved *PI3K* inhibitor). High rates of *MAP3K1* SNVs were identified, however, none of these have been reported as clinically actionable and further studies will be needed to evaluate their significance.

In summary, diagnostic panels such as MammaSeqTM can provide important genomic alteration information that can be used to evaluate BC patient disease heterogeneity and identify actionable drug targets.

ACKNOWLEDGEMENTS

This project benefited from resources provided by the University of Pittsburgh HSCRF Genomics Research Core and the University of Pittsburgh Center for Research Computing. The authors would like to thank Fangping Mu for bioinformatics assistance. This work was supported by funding from Selçuk University.

REFERENCES

1. Stanta, G. & Bonin, S. Overview on Clinical Relevance of Intra-Tumor Heterogeneity. *Front. Med.* (2018). doi:10.3389/fmed.2018.00085
2. Almendro, V. *et al.* Genetic and phenotypic diversity in breast tumor metastases. *Cancer Res.* (2014). doi:10.1158/0008-5472.CAN-13-2357-T
3. Smith, N. G. *et al.* Targeted mutation detection in breast cancer using MammaSeq[®]. *Breast Cancer Res.* **21**, 22 (2019).
4. Merino, G. A. *et al.* TarSeqQC: Quality control on targeted sequencing experiments in R. *Hum. Mutat.* (2017). doi:10.1002/humu.23204
5. Thermofisher. Ion Torrent. Available at: <https://ionreporter.thermofisher.com/ionreporter/help/GUID-A6D3AE1A-84C6-4B66-AA0D-69244576D2FE.html>.
6. Masica, D. L. *et al.* CRAVAT 4: Cancer-related analysis of variants toolkit. *Cancer Res.* (2017). doi:10.1158/0008-5472.CAN-17-0338
7. Talevich, E., Shain, A. H., Botton, T. & Bastian, B. C. CNVkit: Genome-Wide Copy Number Detection and Visualization from Targeted DNA Sequencing. *PLoS Comput Biol* **12**, (2016).
8. Huang, L. *et al.* The cancer precision medicine knowledge base for structured clinical-grade mutations and interpretations. *J. Am. Med. Informatics Assoc.* **24**, 513–519 (2016).
9. R Development Core Team. R: A Language and Environment for Statistical Computing. *R Found. Stat. Comput.* (2016). doi:10.1007/978-3-540-74686-7
10. Chen, H. & Boutros, P. C. VennDiagram: A package for the generation of highly-customizable Venn and Euler diagrams in R. *BMC Bioinformatics* (2011). doi:10.1186/1471-2105-12-35
11. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* (2016). doi:10.1093/bioinformatics/btw313
12. Chalhoub, N. & Baker, S. J. PTEN and the PI3-Kinase Pathway in Cancer. *Annu. Rev. Pathol. Mech. Dis.* (2009). doi:10.1146/annurev.pathol.4.110807.092311
13. Pezo, R. C. *et al.* Impact of multi-gene mutational profiling on clinical trial outcomes in metastatic breast cancer. *Breast Cancer Res Treat* **168**, (2018).
14. Lefebvre, C. *et al.* Mutational Profile of Metastatic Breast Cancers: A Retrospective Analysis. *PLoS Med.* (2016). doi:10.1371/journal.pmed.1002201
15. Ciriello, G. *et al.* Comprehensive Molecular Portraits of Invasive Lobular Breast Cancer. *Cell* (2015). doi:10.1016/j.cell.2015.09.033
16. Curtis, C. *et al.* The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* (2012). doi:10.1038/nature10983
17. Pereira, B. *et al.* The somatic mutation profiles of 2,433 breast cancers refines their genomic and transcriptomic landscapes. *Nat. Commun.* (2016). doi:10.1038/ncomms11479
18. FDA. FDA approves first PI3K inhibitor for breast cancer. (2019). Available at: <https://www.fda.gov/news-events/press-announcements/fda-approves-first-pi3k-inhibitor-breast-cancer>.
19. Gasco, M., Shami, S. & Crook, T. The p53 pathway in breast cancer. *Breast Cancer Res.* (2002). doi:10.1186/bcr426
20. Rebbeck, T. R. *et al.* Hormone-dependent effects of FGFR2 and MAP3K1 in breast cancer susceptibility in a population-based sample of post-menopausal African-American and European-American women. *Carcinogenesis* (2009). doi:10.1093/carcin/bgn247
21. Ahmed, M. & Rahman, N. ATM and breast cancer susceptibility. *Oncogene* (2006). doi:10.1038/sj.onc.1209873
22. Tan, M. & Yu, D. Molecular mechanisms of ErbB2-mediated breast cancer chemoresistance. *Advances in Experimental Medicine and Biology* (2007). doi:10.1007/978-0-387-74039-3_9
23. Friedigkeit, N. *et al.* Intrinsic subtype switching and acquired ERBB2/HER2 amplifications and mutations in breast cancer brain metastases. *JAMA Oncol.* (2017).

- doi:10.1001/jamaoncol.2016.5630
24. Maxwell, K. N. *et al.* BRCA locus-specific loss of heterozygosity in germline BRCA1 and BRCA2 carriers. *Nat. Commun.* (2017). doi:10.1038/s41467-017-00388-9
25. van Bragt, M. P. A., Hu, X., Xie, Y. & Li, Z. RUNX1, a transcription factor mutated in breast cancer, controls the fate of ER-positive mammary luminal cells. *Elife* (2014). doi:10.7554/eLife.03881
26. Li, Q. *et al.* RUNX1 promotes tumour metastasis by activating the Wnt/ β -catenin signalling pathway and EMT in colorectal cancer. *J. Exp. Clin. Cancer Res.* (2019). doi:10.1186/s13046-019-1330-9
27. Farabaugh, S. M., Boone, D. N. & Lee, A. V. Role of IGF1R in breast cancer subtypes, stemness, and lineage differentiation. *Frontiers in Endocrinology* (2015). doi:10.3389/fendo.2015.00059
28. Chakravarty D, Gao J, Phillips S, Kundra R, Zhang H, Solit DB, Schultz N. OncoKB: A Precision Oncology Knowledge Base. *JCO Precis Oncol.* 2017.
29. Turner, N. *et al.* FGFR1 amplification drives endocrine therapy resistance and is a therapeutic target in breast cancer. *Cancer Res.* (2010). doi:10.1158/0008-5472.CAN-09-3746
30. Zhu, H. *et al.* IDH1 R132H mutation enhances cell migration by activating AKT-mTOR signaling pathway, but sensitizes cells to 5-FU treatment as NADPH and GSH are reduced. *PLoS One* (2017). doi:10.1371/journal.pone.0169038
31. Leontiadou, H., Galdadas, I., Athanasiou, C. & Cournia, Z. Insights into the mechanism of the PIK3CA E545K activating mutation using MD simulations. *Sci. Rep.* (2018). doi:10.1038/s41598-018-27044-6
32. Dean, L. *Trastuzumab (Herceptin) Therapy and ERBB2 (HER2) Genotype. Medical Genetics Summaries* (2012).
33. Park, S. *et al.* Higher expression of androgen receptor is a significant predictor for better endocrine-responsiveness in estrogen receptor-positive breast cancers. *Breast Cancer Res. Treat.* (2012). doi:10.1007/s10549-011-1950-z
34. Kensler, K. H. *et al.* Prognostic and predictive value of androgen receptor expression in postmenopausal women with estrogen receptor-positive breast cancer: Results from the Breast International Group Trial 1-98. *Breast Cancer Res.* (2019). doi:10.1186/s13058-019-1118-z
35. Ana Sahores, María May, Gonzalo R. Sequeira, Cynthia Fuentes, Britta Jacobsen, Claudia Lanari, C. A. L. Targeting FGFR with BGJ398 in Breast Cancer: Effect on Tumor Growth and Metastasis. *Curr. Cancer Drug Targets* **18**, (2018).
36. Gao, J. *et al.* Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci. Signal.* (2013). doi:10.1126/scisignal.2004088

Characteristics	n=41
AGE (mean (SD))	50.51 (13.08)
PURITY (%)	
HIGH	26 (63.4)
MEDIUM	10 (24.4)
LOW	3 (7.3)
N/A	2 (4.9)
ER STATUS = Positive (%)	24 (58.5)
PR STATUS = Positive (%)	14 (34.1)
HER2 SCORE (%)	
3	24 (58.5)
2	13 (31.7)
1	2 (4.9)
0	2 (4.9)
RADIO THERAPY = Yes (%)	28 (68.3)
VITAL STATUS = Dead (%)	7 (17.1)
METASTASIS TISSUE (%)	
BONE	17 (41.5)
MIXED	9 (22.0)
LUNG	5 (12.2)
LYMPH NODE	5 (12.2)
BRAIN	2 (4.9)
PLEURA	2 (4.9)
LIVER	1 (2.4)
BREAST CANCER SUBTYPE (%)	
IDC	36 (87.8)
ILC	3 (7.3)
OTHER/UNKNOWN	2 (4.9)

Table 1: Patient and specimen characteristics

Sample ID	Gene	Alteration	Allele Frequency	Copy Number	Detected by	PMKB Tier	OncoKB Level
lonXpress_001	ERBB2	Gain	-	7	Both	1	1
lonXpress_008	ERBB2	Gain	-	41	Both	1	1
lonXpress_010	ERBB2	Gain	-	12	Both	1	1
lonXpress_011	ERBB2	Gain	-	13	Both	1	1
lonXpress_012	ERBB2	Gain	-	27	Both	1	1
lonXpress_014	ERBB2	Gain	-	20	Both	1	1
lonXpress_017	ERBB2	Gain	-	9	Both	1	1
lonXpress_018	ERBB2	Gain	-	11	Both	1	1
lonXpress_031	ERBB2	Gain	-	13	Both	1	1
lonXpress_033	ERBB2	Gain	-	29	Both	1	1
lonXpress_036	ERBB2	Gain	-	18	Both	1	1
lonXpress_038	ERBB2	Gain	-	8	PMKB	1	1
lonXpress_039	ERBB2	Gain	-	14	PMKB	1	1
lonXpress_044	ERBB2	Gain	-	7	PMKB	1	1
lonXpress_045	ERBB2	Gain	-	9	Both	1	1
lonXpress_013	PIK3CA	E545K	0.79	-	PMKB	1	3
lonXpress_015	PIK3CA	E545K	0.32	-	Both	1	3
lonXpress_036	PIK3CA	E545K	0.41	-	Both	1	3
lonXpress_001	PIK3CA	H1047R	0.41	-	Both	2	3
lonXpress_010	PIK3CA	H1047R	0.31	-	Both	2	3
lonXpress_011	PIK3CA	H1047R	0.50	-	Both	2	3
lonXpress_020	PIK3CA	H1047R	0.31	-	Both	2	3
lonXpress_021	PIK3CA	H1047R	0.30	-	Both	2	3
lonXpress_029	PIK3CA	H1047R	0.21	-	Both	2	3
lonXpress_037	PIK3CA	H1047R	0.36	-	Both	2	3
lonXpress_043	PIK3CA	H1047R	0.49	-	Both	2	3
lonXpress_037	PIK3CA	R88Q	0.35	-	PMKB	-	3
lonXpress_001	AR	Gain	-	9	Both	1	-
lonXpress_001	TP53	E204*	0.54	-	PMKB	2	-
lonXpress_046	FGFR1	Gain	-	8	Both	2	-
lonXpress_037	IDH1	R132H	0.36	-	Both	2	-
lonXpress_034	ATM	F858L	0.70	-	Both	3	-
lonXpress_040	JAK3	V722I	0.53	-	OncoKB	3	-

Table 2: Clinically significant variants identified and annotated by OncoKB and PMKB along with information on allele frequencies and copy number changes.

Figure 1: Oncoprint showing the genetic alterations identified by the MammaSeq™ gene panel in Turkish BC Cohort. On left the mutational frequency per gene for this cohort is reported. The corresponding genes are labelled on the right side. The barplots beside the gene labels report the mutational count. Bottom color bar represent the annotation based on clinical variables. The topmost barplots represent the mutation count per sample. IDC: Invasive ductal cancer and ILC: Invasive lobular cancer.

Figure 2: Overlap between interpretations from OncoKB and PMKB on top. Significance categories used by these databases on bottom.