Analysis of Circulating Tumor DNA to Predict Neoadjuvant Therapy Effectiveness and Breast Cancer Recurrence

Shuai Hao, Wuguo Tian, Jianjie Zhao, Yi Chen, Xiaohua Zhang, Bo Gao, Yujun He, Donglin Luo

Department of Breast, Thyroid Surgery, Daping Hospital, Army Medical University, Chongqing, China

ABSTRACT

Purpose: Real-time detection and intervention can be used as potential measures to markedly decrease breast cancer mortality. Assessment of circulating tumor DNA (ctDNA) may offer great benefits for the management of breast cancer over time. However, the use of ctDNA to predict the effectiveness of neoadjuvant treatment and recurrence of breast cancer has rarely been studied.

Methods: We prospectively recruited 31 breast cancer patients with 4 subtypes. Three time points were set in this study, including before any therapy (C1), during surgery (T), and six months after surgery (C2). We collected peripheral blood samples from all 31 patients at C1, tumor tissue from all 31 patients at T, and peripheral blood samples from 25 patients at C2. Targeted 727-gene panel sequencing was performed on ctDNA from all blood samples and tissue DNA from all tissue samples. Somatic mutations were detected and analyzed using a reference standard pipeline. Statistical analysis was performed to identify possible associations between ctDNA profiles and clinical outcomes.

Results: In total, we detected 159, 271, and 70 somatic mutations in 30 C1 samples, 31 T samples, and 12 C2 samples, respectively. We identified specific genes, such as PIK3CA, TP53, and KMT2C, which were highly mutated in the tissue samples. Furthermore, mutated KMT2C observed in ctDNA of the C2 samples may be an indicator of breast cancer recurrence.

Conclusion: Our study highlights the potential of ctDNA analysis at different timepoints for assessing tumor progression and treatment effectiveness, as well as prediction of breast cancer recurrence.

Keywords: Breast neoplasms; Circulating tumor DNA; Neoadjuvant therapy; Recurrence

INTRODUCTION

Breast cancer is the most common cancer in Chinese women, with 304,000 new cases and 70,000 mortalities annually [1]. Breast cancer can be classified into 4 subtypes based on estrogen receptor (ER+), human epidermal growth factor 2 (HER2+), and Ki-67 expression [2]. There are obvious differences in the treatment methods and prognosis for different breast cancer subtypes. For example, ER+ patients have better prognosis when treated with endocrine therapy or combination treatments. The treatment options for patients with
THE AUTHORS DECLARE THAT THEY HAVE NO COMPETING INTERESTS.

CONCEPTUALIZATION: Hao S, Tian W, Zhao J, Chen Y, Zhang X, Gao B, He Y, Luo D; DATA CURATION: Tian W, Zhao J, Chen Y, Zhang X, Gao B, He Y; FUNDING ACQUISITION: Luo D; WRITING - REVIEW & EDITING: Hao S, Luo D.

HER2+ or triple-negative breast cancer (TNBC) are limited, resulting in poor prognosis [3]. Since 2015, neoadjuvant chemotherapy (NAC) has become the standard treatment for locally advanced breast cancer [4]. Previous studies have shown that most patients exhibit pathological complete response (pCR) after NAC, while few patients display stable disease (SD) or even progression. In particular, the pCR rates in HER2+ breast cancer and TNBC are better indicators of prognosis than that in hormone receptor-positive (HR+) patients [5].

As a non-invasive detection method, circulating tumor DNA (ctDNA) sequencing has been widely used for early screening, diagnosis, treatment, and prognosis of breast cancer [6-8]. A potential advantage of ctDNA detection and analysis is the near-real-time monitoring of a patient’s response to treatment, which facilitates rapid feedback on whether a given therapy works well, thereby allowing treatment options to be adjusted in a timely manner. However, to date, the value of ctDNA in the evaluation of breast cancer treatment with neoadjuvant chemotherapy remains unclear.

To our knowledge, this study is the first to evaluate mutational characteristics in ctDNA from patients before and after NAC, as well as in tissues during surgery. Based on the mutations detected, we further evaluated the feasibility of using ctDNA as a new biomarker for NAC response prediction. At the same time, we explored whether ctDNA analysis six months after neoadjuvant treatment can predict breast cancer relapse.

METHODS

Sample preparation
In the present study, we prospectively recruited 31 breast cancer patients and collected peripheral blood samples before neoadjuvant therapy (C1) from 31 patients, tissue samples (T) during surgery from 31 patients, and peripheral blood samples 6 months after surgery (C2) from 25 patients. Ten milliliters of peripheral blood was drawn from each patient at C1 and C2 using a Cell-Free DNA BCT (Streck, Inc., La Vista, USA). Meanwhile, widely used clinical cancer biomarkers, such as carbohydrate antigen 15-3 (CA15-3), carbohydrate antigen 12-5 (CA12-5), and carcinoembryonic antigen (CEA), were detected at the same 3 time points. All patients provided informed consent, and the study was approved by the Institutional Review Board (IRB) of Daping Hospital, Army Medical University (IRB No. 2018(57)). The study protocol adhered to the principles of the Declaration of Helsinki. Normative follow-up within 2 years was also performed for all enrolled patients. Regarding the 10 mL of peripheral blood, plasma was separated from the blood sample within 72 hours by centrifugation (1,600 × g at 4°C for 15 minutes). The supernatants were further centrifuged at 16,000 × g at 4°C for 15 minutes. Plasma samples and peripheral blood cells were used to extract cell-free DNA (cfDNA) and genomic DNA (gDNA) according to the respective operating instructions. Tumor DNA (tDNA) was also extracted from tissue samples based on the standard protocol.

Library construction, sequencing, and processing
In this study, a 727-gene panel covering a 2.8 megabase (Mb) region was designed to capture the target DNA fragments. The lower limit of the target panel detection was ≥ 0.50% for somatic single nucleotide variants (SNVs) and small insertions/deletions (InDels) and ≥ 3 copies for copy number variants (CNVs). The list of genes detected in this study is presented in Supplementary Table 1. All of the collected ctDNA from C1 and C2 and tDNA from tissue samples were subjected to DNA library preparation using a Kapa DNA library preparation...
kit (Kapa Biosystems, Wilmington, USA) while the gDNA library was prepared using an Illumina TruSeq DNA library preparation kit (Illumina, San Diego, USA). The DNA libraries were hybridized to our custom-designed 727-gene probes (Nanodigmbio, Nanjing, China). Then, DNA sequencing was conducted using a DNBSEQ-2000 sequencer (BGI, Shenzhen, China). The raw sequencing data was filtered according to the default parameters to retain high quality reads for subsequent analysis. The clean reads were aligned to human genome assembly (hg19) using the Burrows-Wheeler Aligner (BWA-0.7.17-r1188). Removal of duplicate reads, local realignments, and base-quality recalibrations were performed using the Genome Analysis Toolkit (GATK 4.0.12.0). The resulting binary alignment map (BAM) files were used for subsequent variant calling analysis.

**Somatic mutation calling**

SNVs/small InDels were detected using GATK (4.0.12.0) according to the reference standard pipeline [9]. Germline variants were detected using the HaplotypeCaller in GATK with the default parameters [10]. For all mutational analyses, matched gDNA for each sample was used as the matched control. In short, the peripheral blood sample before neoadjuvant therapy, tissue sample during surgery, and peripheral blood sample six months after surgery were compared to the matched normal samples to exclude germline variants. A Panel of Normal (PoN) was generated from the normal samples to improve the variant calling results. Unreliable somatic mutations found in the PoN, Single Nucleotide Polymorphism Database, or 1000 Genomes Project database were filtered out, resulting in high quality among the remaining detected SNVs and small InDels [11]. In order to verify the detected variants, manual inspection of the BAM files was conducted using the Integrative Genomics Viewer. Tumor mutational burden (TMB) was defined as the number of somatic mutations per Mb.

**CNV calling**

CNVs were estimated by paired samples with Varscan2 according to the default parameters [12] as follows: first, mpileup was performed on the BAM files of normal and tumor samples using samtools mpileup, and guanine-cytosine correction was performed on the mpileup results with CopyCaller software. Second, the regions with different copy numbers were cut with the circular binary segmentation algorithm. Finally, CNVs were obtained after combining the segments of candidate CNVs. As in somatic mutation calling, a CNV PoN from the normal samples was created to improve the results.

**Statistical analysis**

Univariate and multivariate Cox proportional hazard analyses were conducted to assess the correlation between clinical features and the treatment outcome (pCR) of neoadjuvant therapy. A Kaplan-Meier plot was also used to identify the correlation between ctDNA samples from the C2 time point and tumor recurrence. A $\chi^2$ test was utilized to determine whether the level of ctDNA in the blood before neoadjuvant treatment was related to its therapeutic efficacy. Statistical significance was defined as a 2-sided p-value of < 0.05. Statistical analyses were performed using SAS 9.4 software (SAS Institute, Cary, USA).

**RESULTS**

**Sample cohort description**

In total, 31 breast cancer patients (mean age: 48.2 ± 7.0 years) were recruited in this study: 17 cases with the luminal subtype (54.83%), 7 cases with HER2 overexpression (22.58%), and
7 cases with TNBC (22.58%) (Table 1 and Supplementary Table 2). It is worth noting that, for various reasons, 6 blood samples were not obtained at the 6-month postoperative time node (P6C2, P7C2, P17C2, P21C2, P24C2, and P30C2). Among all patients, 15 (48.39%) were diagnosed with left breast cancer while the remaining ones had right breast cancer. The body mass indexes (BMIs) of 17 patients (54.84%) were within normal range (18.5–24 kg/m²), and the remaining 14 patients were overweight (BMI > 24 kg/m²). Twenty-three patients (74.19%) received taxotere, anthracycline, and cyclophosphamide (TAC) while the remaining eight (25.81%) received taxotere, carboplatin, and herceptin (TCH). Interestingly, the pathological subtype following needle biopsy in three patients (P2, P17, and P28) changed after breast

| Characteristics | Patients (n=31) |
|-----------------|----------------|
| Age (yr)        | 48.2 ± 7.0     |
| TNM stage       |                |
| IIB             | 10 (32.2)      |
| IIIA            | 18 (58.1)      |
| IIB             | 2 (6.5)        |
| IIIA            | 1 (3.2)        |
| Molecular subtype |            |
| Luminal A       | 9 (29.0)       |
| Luminal B       | 8 (25.8)       |
| HER2+++         | 7 (22.6)       |
| TNBC            | 7 (22.6)       |
| Lateral of breast cancer |     |
| Left            | 15 (48.4)      |
| Right           | 16 (51.6)      |
| NAC             |                |
| TAC             | 23 (74.2)      |
| TCH             | 8 (25.8)       |
| BMI (kg/m²)     |                |
| < 18.5          | 0 (0.0)        |
| ≥ 18.5 and < 24 | 17 (54.8)      |
| ≥ 24            | 14 (45.2)      |
| RECIST          |                |
| CR              | 1 (3.2)        |
| PR              | 19 (61.3)      |
| SD              | 11 (35.5)      |
| LN              |                |
| 0               | 13 (41.9)      |
| 1–3             | 9 (29.0)       |
| > 3             | 9 (29.0)       |
| MP              |                |
| 1               | 3 (9.7)        |
| 2               | 10 (32.3)      |
| 3               | 13 (41.9)      |
| 4               | 1 (3.2)        |
| 5               | 4 (12.9)       |
| Surgery         |                |
| MRM             | 28 (90.3)      |
| Mastectomy + SLNB | 2 (6.5)  |
| Cancer embolus  |                |
| Yes             | 9 (29.0)       |
| No              | 22 (71.0)      |

Values are expressed as mean ± standard deviation or number (%).

TNM = tumor, node, metastasis; HER2 = human epidermal growth factor 2; TNBC = triple-negative breast cancer; NAC = neoadjuvant chemotherapy; TAC = taxotere, anthracycline, and cyclophosphamide; TCH = taxotere, carboplatin, and herceptin; BMI = body mass index; RECIST = Response Evaluation Criteria in Solid Tumors; CR = complete response; PR = partial response; SD = stable disease; MP = Miller-Payne; MRM = modified radical mastectomy; SLNB = sentinel lymph node biopsy.
resection. Specifically, patient P2 changed from HER2 overexpression to Luminal A. Patient P17 changed from Luminal B to Luminal A with a dramatic decrease in Ki-67 expression from 40.00% to 3.00%. Patient P28 changed from TNBC to Luminal B. These results suggest that the heterogeneity of multiple regions is common in breast cancer. Regarding the widely used clinical cancer biomarkers, 3 out of 31 patients (P5C1, P16C1, and P17C1) displayed abnormal results at time point C1, 1 (P5T1) at T, and another 1 (P5C2) at C2. In patient P5, the CA 15-3 level was 37.05 U/mL and the CEA level was 28.41 ng/mL at C1 (the normal ranges for CA15-3 and CEA are 0–28 U/mL and 0–5 ng/mL, respectively). Even after NAC and surgery, these conditions did not improve, and the patient relapsed 9 months after time point T (CA15-3: 44.27 U/mL and CEA: 15.11 ng/mL).

Target-capture sequencing and genetic profiles
For each patient, we prospectively collected peripheral blood samples before administering any therapy (C1 as the baseline), tissue samples during breast dissection (T), and blood samples at six months after breast dissection (C2). A custom panel was applied to achieve target-capture sequencing for all samples (87 samples from 31 patients). For the raw sequencing data, the average sequencing capture efficiency for the C1, T, and C2 samples was 72.89% (64.20–78.72%), 74.98% (59.72%–80.62%), and 72.80% (67.53%–81.05%), respectively, and the effective target sequencing depth for the C1, T, and C2 samples was 1,010.19X (633.1X–1,650.01X), 1071.61X (474.41X–1,624.33X), and 925.15X (625.15X–1,383.41X), respectively (Supplementary Table 3). Overall, 159, 271, and 70 somatic mutations were identified in the C1, T, and C2 samples, respectively. Somatic mutations were detected in 30 out of 31 C1 samples and 12 out of 25 C2 samples (Supplementary Table 4). Compared to the T samples, for which a mean of 8.7 (2–23) mutations were detected per sample, the number of mutations per C1 sample was significantly lower (5.1 [1–14]). When comparing somatic mutations between the C1 and T samples, we found that 46.54% of these mutations at C1 were also identified in T samples. In addition, 10, 26, and 10 somatic hotspot mutations were identified in the C1, T, and C2 samples, respectively. Among them, somatic hotspot mutations were detected in 9 out of 31 (29.03%) C1 samples and 8 out of 25 (32.0%) C2 samples. When comparing somatic hotspot mutations between the C1 and T samples, we found that 80.0% of these mutations at C1 were also identified in T samples. Interestingly, these somatic hotspot mutations belong to one kind, namely KMT2C c.5053G>T. These results revealed high heterogeneity between blood and tissue samples in the early and middle stages of our study [13].

Among the tissue samples, the most common somatically mutated genes were PIK3CA (48.50%), TP53 (41.90%), and KMT2C (9.70%). However, when considering all 87 samples, the most common mutated genes were KMT2C (20.69%), PIK3CA (18.39%), and TP53 (18.39%) (Figure 1). These discrepancies clearly contributed to the heterogeneity of samples from different time points. In addition to somatic mutations, we also identified 6 germline mutations (19.35%) in the 31 breast cancer patients involving BRCA2, CHEK2, and MUTYH. They were BRCA2 c.5645C>A (P1), BRCA2 c.987_988insA (P7), MUTYH c. 892-2A>G (P15), BRCA2 c.31del (P17), CHEK2 c.1651dup (P29), and BRCA2 c.3940_3941del (P31) (Supplementary Table 5). Through univariate and multivariate analyses, these mutations were identified to be statistically unrelated to pCR and recurrence. However, it was very interesting that five out of the six identified germline mutations occurred in the luminal subtype of breast cancer. No recurrent BRCA1/2 mutations were found, indicating that no founder mutations were present in the Chinese population [14].
We detected mutations in eight C1 samples, three T samples, and seven C2 samples. Overall, 23 out of 87 (26.44%) samples were detected to have abnormal gene copy numbers; among them, 19 out of 23 (82.61%) were from tissue samples while the remaining 4 (17.29%) were from peripheral blood samples (Supplementary Table 6). A total of 131 CNVs were found with 107 detected in gene amplification and the remaining 24 in deletion. In the tissue samples, 35 genes had CNVs, and the most common ones were HRAS amplification (14 out of 119 [11.76%]), CND1 amplification (13 out of 119 [10.92%]), AKT1 amplification (12 out of 119 [10.08%]), NOTCH1 amplification (11 out of 119 [9.24%]), and TSC2 amplification (9 out of 119 [7.56%]). Among the relapsed patients, no CNVs were found in C2 samples. Based on these results, CNVs may not be a good biomarker for early breast cancer recurrence monitoring.

**ctDNA KMT2C c.5053G>T association with breast cancer relapse**

In clinical contexts, cancer embolus and number of lymph nodes serve as high risk indexes for breast cancer patients with poor prognosis. Univariate analysis found that patients with high lymph node count (≥ 4) and SD after NAC relapsed or progressed within a short period (p = 0.012 and p = 0.036, respectively) while cancer embolus did not correlate with the period of relapse (p > 0.05). We thus concluded that lymph node status and pCR could be used as predictors of breast cancer recurrence.

To investigate whether postoperative blood ctDNA could also be used as a predictor of recurrence after neoadjuvant therapy, we compared somatic mutations from the C1, T, and C2 samples. We detected KMT2C mutations in eight C1 samples, three T samples, and seven C2 samples. Then, we further analyzed the seven patients with relapsed breast cancer. Two patients failed to produce C2 samples (P6C2 and P24C2) so KMT2C mutation status was unknown, one patient...
(P10C2) was detected with no mutation, and four patients (P5C2, P16C2, P18C2, and P27C2) were all detected to have \( \text{KTM2C} \ c.5053G>T \). Importantly, survival analysis (Kaplan-Meier plot) showed that ctDNA \( \text{KTM2C} \ c.5053G>T \) was significantly associated with postoperative breast cancer recurrence \( (p = 0.0014) \) (Figure 2). Although the allelic frequency of \( \text{KMT2C} \ c.5053G>T \) dropped in P5 and P18, it was still relatively high (Figure 3A-D). These findings imply that ctDNA mutations could be used as a potential predictor of breast cancer recurrence after neoadjuvant therapy. As for the CA 15-3, CA 12-5, and CEA results, only one relapsed patient (P5) displayed abnormal levels while the remaining six were normal, suggesting that traditional tumor biomarkers could not meet the requirements for early tumor recurrence detection.

Figure 2. Kaplan-Meier plot of \( \text{KMT2C} \) mutation and relapse time (month). FPS = French prognostic score.

Figure 3. Mutated allele frequency of \( \text{KMT2C} \ c.5053G>T \) in 7 patients at different timepoints. The patients P5, P4 and P31 were identified \( \text{KMT2C} \) mutation in tissue samples, while P16, P18, P27, P22 samples were undetectable in tissue samples.
Finally, we analyzed TMB and mutations derived from the C1 blood samples. Of the 8 KMT2C mutations identified in C1 samples, 4 patients displayed SD after neoadjuvant therapy while the remaining 4 had a pCR. The association between KMT2C and neoadjuvant therapy efficiency was insignificant \( (p > 0.05) \) based on a \( \chi^2 \) test. The effectiveness of TMB was also insignificant. However, we found that some of the SD patients with or without KMT2C mutations had relatively high TMB, including patients P6 (7 mutations), P15 (6 mutations), and P23 (7 mutations), which suggests that the effectiveness of neoadjuvant therapy may be predictable by KMT2C combined with high TMB in large cohorts.

**DISCUSSION**

Here, we performed custom-panel high-depth sequencing on baseline blood, tissue, and postoperative blood samples from 31 newly diagnosed breast cancer patients and detected 159 (30 out of 31 samples), 271 (31 out of 31 samples), and 70 (12 out of 25 samples) somatic mutations, respectively (Supplementary Table 4). Among them, 46.54\% of mutations in the baseline blood samples were also detected in the tissue samples, suggesting that the heterogeneity of tissue and blood affects composition. Regarding the 6-month postoperative blood samples, 57.14\% of observed mutations were also present in the baseline blood samples, suggesting that blood can better overcome the heterogeneity problem. Interestingly, mutations in the blood at 6 month post-operation were detected in only 31.42\% of the tissue samples, meaning that changes in the postoperative blood ctDNA may be associated with treatment. The three most common somatically mutated genes were PIK3CA, TP53, and KMT2C (Figure 1). Previous studies have reported that the mutation rate of PIK3CA in HR+ breast cancer patients ranged from 40\% to 70\% and is related to tumor growth, resistance to endocrine therapy, and poor overall prognosis [15,16]. Similarly, in this study, the PIK3CA mutation rate was 48.5\%. In 2019, the United States Food and Drug Administration approved alpelisib in combination with fulvestrant for treatment in postmenopausal women and men with HR+, HER2−, or PIK3CA− mutated advanced or metastatic breast cancer[17,18], suggesting that the PIK3CA mutation was of great significance in breast cancer patients, especially for the treatment of HR+ breast cancer after metastasis or progression. We also found that TP53 mutations were mainly detected in tissue samples from the seven TNBC patients. This finding about the high mutation frequency of TP53 in basal-like breast cancers was similar to a reported study for HGS-OvCa with TP53 mutation patterns in almost 100\% of samples [19]. KMT2C, which encodes an H3K4 histone methyltransferase, was reported as a driver gene for metastatic breast cancer that expressed hormone receptors but not HER2 at high levels [20]. KMT2C has been previously reported with mutation rates of 6.45\%, 8.16\%, and 1.96\% in The Cancer Genome Atlas, Molecular Taxonomy of Breast Cancer International Consortium, and French studies, respectively [21-24]. In this study, the mutation rate of KMT2C was 20.69\% in all 87 samples and 9.67\% (3 out of 31) in tissue samples. This result was consistent with that of a study from Singapore and Korea [25]. The 20 top mutated genes also included RECQL4 (11.49\%), ERBB2 (10.34\%), and RAD51B (9.21\%). Interestingly, a recent study showed that the RecQ family is involved in DNA mismatch repair, indicating its role in cancer pathogenesis and development [26].

In addition to somatic mutations, we also discovered 6 germline mutations involving BRCA2, CHEK2, and MUTHY in the 31 patients, but these mutations were statistically unrelated to pCR and recurrence. In reports by previous researchers, CHEK2 was found to be an important moderate-penetrance breast cancer predisposition gene. This gene functions as a responder primarily to double-strand break DNA damage repair. In a breast cancer study involving 7,657 BRCA1/2-negative breast cancer patients, the mutation rate of CHEK2 was 0.34\% [27]. For the
MUTYH gene, a large number of studies have reported its association with an increased risk of colorectal cancer, but its correlation with breast cancer remains unclear [28]. Through MUTYH c.892-2A>G was reported in a Chinese population with a mutation rate of 5.8%, the rate in East Asian healthy individuals was 2.77%. These discrepancies between our results and those of previous research may be attributed to small sample sizes.

Our study detected that four out of five relapsed breast cancer patients carried the KMT2C c.5053G>T mutation. Kaplan-Meier survival and multivariate analyses found that KMT2C, like other clinical factors, was significantly associated with postoperative breast cancer recurrence. It is interesting that previous studies have identified KMT2C as a driving factor of HR+ breast cancer [20,21]. In this study, three out of the four relapsed patients had TNBC, and one had HER2+. A study from Thailand showed that KMT2C has a higher mutation rate in TNBC, which suggests that KMT2C mutations may be related to the occurrence and development of breast cancer in Asian patients [29]. For relapsed patient P10C2, no KMT2C mutation was detected in baseline blood, tissue, and 6-month postoperative blood samples. However, the patient carried the TP53 mutation, and their TMB was high with SD present even after neoadjuvant therapy. We speculate that these factors may lead to recurrence in TNBC patients.

We also identified that 3 patients (P4, P22, and P31) carried the KMT2C c.5053G>T mutation without relapse at the time of follow-up (Figure 3E and F). Patient P4 was diagnosed with right HR+ breast cancer at 40 years old and treated with TAC. Following NAC, she displayed a pCR, which indicated a good prognosis. Patients P22 and P31 were also diagnosed with the HR+ subtype and were identified as SD and pCR after neoadjuvant therapy, respectively. Though they each carried the KMT2C mutation, it had a decreased mutated allele frequency. Furthermore, considering the short follow-up time, their future recurrent statuses remained undetermined. Thus, constant monitoring and follow-up for these patients are required.

To our knowledge, this study is the first to explore the relationship between prognosis and ctDNA after neoadjuvant therapy in Chinese breast cancer patients. We found that the incidence of KMT2C mutations is higher in breast cancer patients, especially among TNBC patients. Kaplan-Meier survival and multivariate analyses indicated that ctDNA KMT2C c.5053G>T could be a predictor of the relapse of breast cancer, while the association between KMT2C and efficiency of neoadjuvant therapy was insignificant (p > 0.05). Unfortunately here, neither CNV nor TMB was a qualified indicator for neoadjuvant therapy and breast cancer recurrence. However, our study has considerable limitations. First, the follow-up period of 2 years was not enough to identify a clinically meaningful association between ctDNA mutations and clinical outcomes. Further studies are certainly necessary. Second, because the tissue samples were formalin-fixed, paraffin-embedded samples, the extracted gDNA may have been degraded, resulting in variant call discrepancies despite utilization of filtering and correction methods during the analysis. Finally, our sample size was limited. There may be biases that do not represent the true characteristics of breast cancer subtypes due to the small study cohort [30].

ACKNOWLEDGMENTS

We thank all researchers for their contributions and all patients for their participation. We also thank Yujian Shi, Yangming Wu, Luhua Lin, Ang Li, Yue Yang, Qihuan Zhi, Qingqing He and Ting Ouyang from Top Gene Tech Co., Ltd. (Guangzhou, China) for next-generation sequencing technical support and scientific suggestions.
SUPPLEMENTARY MATERIALS

Supplementary Table 1
The list of genes detected in this study
Click here to view

Supplementary Table 2
Clinical characteristics of 31 breast cancer patients
Click here to view

Supplementary Table 3
The results of the target-panel sequencing quality control in this study
Click here to view

Supplementary Table 4
Identified somatic mutations from 87 samples
Click here to view

Supplementary Table 5
Pathogenic germline alterations identified in 31 breast cancer patients
Click here to view

Supplementary Table 6
Identified copy number variants from 87 samples
Click here to view

REFERENCES

1. Zheng RS, Sun KX, Zhang SW, Zeng HM, Zou XN, Chen R, et al. Report of cancer epidemiology in China, 2015. Zhonghua Zhong Liu Za Zhi 2019;41:i9-28.
   PUBMED | CROSSREF

2. Dai X, Li T, Bai Z, Yang Y, Liu X, Zhan J, et al. Breast cancer intrinsic subtype classification, clinical use and future trends. Am J Cancer Res 2015;5:2929-43.
   PUBMED

3. Prat A, Fan C, Fernández A, Hoadley KA, Martinello R, Vidal M, et al. Response and survival of breast cancer intrinsic subtypes following multi-agent neoadjuvant chemotherapy. BMC Med 2015;13:303.
   PUBMED | CROSSREF

4. Haddad TC, Goetz MP. Landscape of neoadjuvant therapy for breast cancer. Ann Surg Oncol 2015;22:1408-15.
   PUBMED | CROSSREF

5. Esserman LJ, Berry DA, DeMichele A, Carey L, Davis SE, Buxton M, et al. Pathologic complete response predicts recurrence-free survival more effectively by cancer subset: results from the I-SPY1 TRIAL--CALGB 150007/150012, ACRIN 6657. J Clin Oncol 2012;30:3242-9.
   PUBMED | CROSSREF
6. Buono G, Gerratana L, Bulfoni M, Provinciali N, Basile D, Giuliano M, et al. Circulating tumor DNA analysis in breast cancer: is it ready for prime-time? Cancer Treat Rev 2019;73:73-83.
PUBMED | CROSSREF

7. Zhang X, Zhao W, Wei W, You Z, Ou X, Sun M, et al. Parallel analyses of somatic mutations in plasma circulating tumor DNA (ctDNA) and matched tumor tissues in early-stage breast cancer. Clin Cancer Res 2019;25:6546-53.
PUBMED | CROSSREF

8. Coombes RC, Page K, Salari R, Hastings RK, Armstrong A, Ahmed S, et al. Personalized detection of circulating tumor DNA antedates breast cancer metastatic recurrence. Clin Cancer Res 2019;25:4255-63.
PUBMED | CROSSREF

9. do Valle IF, Giampieri E, Simonetti G, Padella A, Manfrini M, Ferrari A, et al. Optimized pipeline of MuTect and GATK tools to improve the detection of somatic single nucleotide polymorphisms in whole-exome sequencing data. BMC Bioinformatics 2016;17:341.
PUBMED | CROSSREF

10. Ren S, Bertels K, Al-Ars Z. Efficient acceleration of the pair-HMMs forward algorithm for GATK HaplotypeCaller on graphics processing units. Evol Bioinform Online 2018;14:1176934318760543.
PUBMED | CROSSREF

11. Sun P, Zhong Z, Lu Q, Li M, Chao X, Chen D, et al. Mucinous carcinoma with micropapillary features is morphologically, clinically and genetically distinct from pure mucinous carcinoma of breast. Mod Pathol. Epub 2020 May 1. https://doi.org/10.1038/s41379-020-0554-8.
PUBMED | CROSSREF

12. Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, Lin L, et al. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Res 2012;22:568-76.
PUBMED | CROSSREF

13. Shatsky R, Parker BA, Bui NQ, Helsten T, Schwab RB, Boles SG, et al. Next-generation sequencing of tissue and circulating tumor DNA: the UC San Diego Moores Center for personalized cancer therapy experience with breast malignancies. Mol Cancer Ther 2019;18:568-76.
PUBMED | CROSSREF

14. Bhaskaran SP, Chandratre K, Gupta H, Zhang L, Wang X, Cui J, et al. Germline variation in BRCA1/2 is highly ethnic-specific: evidence from over 30,000 Chinese hereditary breast and ovarian cancer patients. Int J Cancer 2019;145:962-73.
PUBMED | CROSSREF

15. Oshiro C, Kagara N, Naoi Y, Shimoda M, Shimomura A, Maruyama N, et al. PIK3CA mutations in serum DNA are predictive of recurrence in primary breast cancer patients. Breast Cancer Res Treat 2015;150:299-307.
PUBMED | CROSSREF

16. Loibl S, von Minckwitz G, Schneeweiss A, Paepke S, Lehmann A, Rezai M, et al. PIK3CA mutations are associated with lower rates of pathologic complete response to anti-human epidermal growth factor receptor 2 (HER2) therapy in primary HER2-overexpressing breast cancer. J Clin Oncol 2014;32:3212-20.
PUBMED | CROSSREF

17. Food and Drug Administration. FDA approves alpelisib for metastatic breast cancer. https://www.fda.gov/drugs/resources-information-approved-drugs/fda-approves-alpelisib-metastatic-breast-cancer. Accessed January 9th, 2020.
PUBMED | CROSSREF

18. Copur MS, Jonglertham P. Alpelisib for PIK3CA-mutated advanced breast cancer. N Engl J Med 2019;381:686-7.
PUBMED | CROSSREF

19. Silwal-Pandit L, Langerød A, Børresen-Dale AL. TP53 mutations in breast and ovarian cancer. Cold Spring Harb Perspect Med 2017;7:a026252.
PUBMED | CROSSREF

20. Bertucci F, Ng CK, Patoursis A, Drioin N, Piscuoglio S, Carluccia N, et al. Genomic characterization of metastatic breast cancers. Nature 2019;569:560-4.
PUBMED | CROSSREF

21. Gala K, Li Q, Sinha A, Razavi P, Dorso M, Sanchez-Vega F, et al. KMT2C mediates the estrogen dependence of breast cancer through regulation of ERα enhancer function. Oncogene 2018;37:4692-710.
PUBMED | CROSSREF

22. Kandoth C, McLellan MD, Vandin F, Ye K, Niu B, Lu C, et al. Mutational landscape and significance across 12 major cancer types. Nature 2013;502:333-9.
PUBMED | CROSSREF

23. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. Nature 2012;490:61-70.
PUBMED | CROSSREF
24. Pereira B, Chin SF, Rueda OM, Vollan HK, Provenzano E, Bardwell HA, et al. The somatic mutation profiles of 2,433 breast cancers refines their genomic and transcriptomic landscapes. Nat Commun 2016;7:11479.

25. Yap YS, Singh AP, Lim JH, Ahn JH, Jung KH, Kim J, et al. Elucidating therapeutic molecular targets in premenopausal Asian women with recurrent breast cancers. NPJ Breast Cancer 2018;4:19.

26. Peng J, Tang L, Cai M, Chen H, Wong I, Zhang P. RECQL5 plays an essential role in maintaining genome stability and viability of triple-negative breast cancer cells. Cancer Med 2019;8:4743-52.

27. Fan Z, Ouyang T, Li J, Wang T, Fan Z, Fan T, et al. Identification and analysis of CHEK2 germline mutations in Chinese BRCA1/2-negative breast cancer patients. Breast Cancer Res Treat 2018;169:59-67.

28. Jian W, Shao K, Qin Q, Wang X, Song S, Wang X. Clinical and genetic characterization of hereditary breast cancer in a Chinese population. Hered Cancer Clin Pract 2017;15:19.

29. Niyomnaitham S, Parinyanitikul N, Roothumnong E, Jinda W, Samarnthai N, Atikankul T, et al. Tumor mutational profile of triple negative breast cancer patients in Thailand revealed distinctive genetic alteration in chromatin remodeling gene. PeerJ 2019;7:e6501.

30. Weiss NS. Cohort studies of the efficacy of screening for cancer. Epidemiology 2015;26:362-4.