

# Serial Circulating Tumor DNA Identification Associated with the Efficacy and Prognosis of Neoadjuvant Chemotherapy in Breast Cancer

**Yidong Zhou**

Peking Union Medical College Hospital

**Yaping Xu**

Geneplus-Beijing

**Changjun Wang**

Peking Union Medical College Hospital

**Yuhua Gong**

Geneplus-Beijing

**Yanyan Zhang**

Geneplus-Beijing

**Ru Yao**

Peking Union Medical College Hospital

**Peng Li**

Peking Union Medical College Hospital

**Xiuli Zhu**

Geneplus-Beijing

**Jing Bai**

Geneplus-Beijing

**Yanfang Guan**

Geneplus-Beijing

**Xuefeng Xia**

Geneplus\_Beijing

**Ling Yang**

Geneplus-Beijing

**Xin Yi**

Geneplus-Beijing

**Qiang Sun** (✉ [sunqiangpumch@sina.com](mailto:sunqiangpumch@sina.com))

Peking Union Medical College Hospital

---

## Research Article

**Keywords:** Circulating tumor DNA, locally advanced breast cancer, neoadjuvant chemotherapy, efficacy, prognosis.

**Posted Date:** February 17th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-220659/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

---

# Abstract

**Background:** Circulating tumor DNA (ctDNA) provides a promising noninvasive alternative to evaluate the efficacy of neoadjuvant chemotherapy (NCT) in breast cancer.

**Methods:** Herein, we collected 63 tissue (aspiration biopsies and resected tissues) and 206 blood samples (baseline, during chemotherapy (Chemo), after chemotherapy (Post-Chemo), after operation (Post-Op), during follow-up) from 32 patients, and performed targeted deep sequencing with a customized 1,021-gene panel.

**Results:** As the results, TP53 (43.8%) and PIK3CA (40.6%) were the most common mutant genes in the primary tumors. At least one tumor-derived mutation was detected in the following number of blood samples: 21, baseline; 3, Chemo; 9, Post-Chemo; and 5, Post-Op. Four patients with pathologic complete response had no tissue mutation in Chemo and Post-Chemo blood. Compared to patients with mutation-positive Chemo or Post-Chemo blood, the counterparts showed a superior primary tumor decrease (median, 86.5% versus 54.6%) and lymph involvement (median, one versus 3.5). All five patients with mutation-positive Post-Op developed distant metastases during follow-up, and the sensitivity of detecting clinically relapsed patients was 71.4% (5/7). The median DFS was 9.8 months for patients with mutation-positive Post-Op but not reached for the others (HR 23.53; 95% CI, 1.904–290.9;  $p < 0.0001$ ).

**Conclusions:** Our study shows that sequential monitoring of blood ctDNA was an effective method for evaluating NCT efficacy and patient survival. Integrating ctDNA profiling into the management of LABC patients might improve clinical outcome.

**Trial registration:** This prospective study recruited LABC patients at Peking Union Medical College Hospital (ClinicalTrials.gov Identifier: NCT02797652).

## Introduction

Breast cancer is one of the most common malignancies for women. According to 2018 global cancer statistics, over 2 million patients were diagnosed as having breast cancer and 630,000 died of the disease worldwide[1]. Currently, neoadjuvant chemotherapy (NCT) is a useful treatment for locally advanced breast cancer (LABC) to downstage and render inoperable disease operable[2]. In addition, NCT can reduce the spatial size of tumors too large for a lumpectomy, thus providing an opportunity to conserve the breast and greatly improve the post-operative quality of life of patients[2]. However, the efficacy evaluation of NCT has not yet been well developed. The clinical imaging methods, such as ultrasound and magnetic resonance imaging, are not always in accordance with pathological assessment[3]. Furthermore, conventional tumor markers, such as CEA and CA15-3, have not shown a strong correlation with response to NCT[4–6]. Because distant relapse is the major cause of cancer death[7, 8], more effective diagnostic tools that stratify patients by distant recurrence risk would be of great clinical significance to deliver personalized healthcare. Overall, the development of novel markers with high prognostic and predictive performance for post-operative relapse risk are crucial to further facilitate clinical decision-making.

Circulating tumor DNA (ctDNA) is derived from the apoptosis, necrosis and secretion of tumor cells and can harbor tumor-specific genetic or epigenetic alterations[9, 10]. CtDNA can be detected in various types of solid tumors, including breast cancer[11]. Several studies have reported that mutant ctDNA is correlated with the response to systemic treatments in breast cancer[3, 12–16]. However, most of these studies only focused on limited epigenetic or genetic markers, such as *RASSF1A* methylation[3] and *TP53* variants[13], and breast cancer is a highly genetically heterogeneous malignancy[17–20]. Such strategies restricted the population that can potentially benefit from ctDNA profiling. In addition, tracking of mutant ctDNA post-operatively could be a useful method for predicting and monitoring distant relapse in breast cancer[12].

In this study, we enrolled 32 of patients with locally advanced and metastatic breast cancer who were eligible for NCT and underwent standardized clinical managements. We collected tumor tissues to profile patient-specific mutation pools using targeted sequencing with a well-designed capture panel and traced them in sequential blood ctDNA around NCT and during post-operative follow-up. Through these efforts, we aimed to explore the utility of ctDNA to evaluate the efficacy and prognosis of neoadjuvant chemotherapy for breast cancer.

## Results

### Experimental cohort and genomic profiling of tissue

An overview of patient enrollment and study design is shown in *Fig. S1*. We initially enrolled 65 patients with locally advanced or metastatic breast cancer between May 2016 and March 2017. After further clinical assessment, 23 patients were excluded because they were ineligible for neoadjuvant chemotherapy. In addition, one patient was excluded due to insufficient tumor tissue for genomic analysis and seven patients were excluded due to drop-out during follow-up. Ultimately, 32 patients were included in the study, and 63 tissue and 206 blood samples were collected (*Fig. S2*). Baseline (pre-treatment) patient characteristics and clinical staging information are shown in Table 1. The median diagnostic age was 47 years (27–67). Twelve patients (37.5%) were menopausal. Most tumors (71.9%, 23/32) were moderately differentiated. Twenty-four (75.0%) patients were estrogen receptor (ER) positive and 21 (65.6%) were progesterone receptor (PR) positive. HER2 overexpression was identified in 11 patients (34.4%). Only four patients (12.5%) were identified with low Ki67 (< 14%). Except for five patients with metastatic disease (15.6%), 5 (15.6%) had stage IIb and 22 (68.8%) patients had stage I disease.

Table 1  
Clinicopathological characteristics and ctDNA detectability  
of analyzed patients

Characteristics	Total = 32
Age at diagnosis, years, median (range)	47 (27 to67)
Menstruation	
Menopausal, n (%)	12 (37.5%)
Premenopausal, n (%)	20 (62.5%)
Pathological examination of the biopsy	
Moderate differentiation, n (%)	23 (71.9%)
Poor differentiation, n (%)	9 (28.1%)
ER status	
Positive, n (%)	24 (75.0%)
Negative, n (%)	8 (25.0%)
PR status	
Positive, n (%)	21 (65.6%)
Negative, n (%)	11 (34.4%)
HER2 Overexpression	
Positive, n (%)	11 (34.4%)
Negative, n (%)	21 (65.6%)
Ki67	
< 14%, n (%)	4 (12.5%)
≥ 14%, n (%)	28 (87.5%)
Luminal typing	
Luminal A, n (%)	4 (12.5%)
Luminal B, n (%)	20 (62.5%)
HER2 overexpression, n (%)	3 (9.4%)
TNBC, n (%)	5 (15.6%)
T stage	
T1, n (%)	4 (12.5%)
T2, n (%)	22 (68.8%)
T3, n (%)	6 (18.7%)
N stage	
N1, n (%)	9 (28.1%)
N2, n (%)	17 (53.2%)
N3, n (%)	6 (18.7%)
M stage	

Characteristics	Total = 32
M0, n (%)	27 (84.4%)
M1, n (%)	5 (15.6%)
Clinical stage	
IIb, n (%)	5 (15.6%)
III, n (%)	22 (68.8%)
IV, n (%)	5 (15.6%)
pCR	
Yes, n (%)	4 (12.5%)
No, n (%)	28 (87.5%)
Baseline ctDNA	
Positive, n (%)	21 (65.6%)
Negative, n (%)	9 (28.1%)
NA*, n (%)	2 (6.3%)
Chemo ctDNA	
Positive, n (%)	3 (9.4%)
Negative, n (%)	29 (90.6%)
Post-Chemo ctDNA	
Positive, n (%)	9 (28.1%)
Negative, n (%)	22 (68.8%)
NA*, n (%)	1 (3.1%)
Post-Op ctDNA	
Positive, n (%)	5 (15.6%)
Negative, n (%)	27 (84.4%)

**Abbreviations:** ER, estrogen receptor; PR, progesterone receptor; TNBC, triple-negative breast cancer; Chemo, chemotherapy; Op, operation.

\*sample unattained.

The genomic status of tumor tissue was obviously changed under the pressure of neoadjuvant chemotherapy. Within 117 and 75 mutations identified in 31 pairs of tissue biopsy and surgically resected tissue samples, only 46 mutations were co-expressed in both samples, and the genomic status was stable in only two patients (P06, P25), indicating a longitudinal heterogeneity induced by chemotherapy (Fig. S3). To eliminate the temporal bias in the evaluation of plasma, we assessed the union set of tissue biopsy and resected tissue mutations in the same patient. As a result, at least one mutation was identified in each patient. In total, 151 mutations were detected with a median of 3 (range, 1–23). *TP53* (n = 14, 43.8%) and *PIK3CA* (n = 13, 40.6%) were the most recurrent mutant genes as expected (Fig. 1A). We also evaluated the concordance between *ERBB2* amplification at baseline and HER2 status determined by immunohistochemical staining and fluorescence in situ hybridization (Fig. 1B). Among 11 patients with HER2-positive tumors, 10 (90.9%) carried *ERBB2* amplification (copy number > 3). All 21 patients with negative HER2 exhibited a normal copy number of *ERBB2*. The overall concordance was 96.9%. Even for the inconsistent one, the copy number of *ERBB2* (2.6) was higher than that of patients with HER2-negative tumors. It did not reach the cut-off value possibly due to the insufficient proportion of tumor cells in tissue sample.

## Clinical Characteristics And Baseline Ctdna

Based on the aforementioned tissue analysis, we tracked so-called tumor-derived mutations in blood ctDNA, including SNVs, Indels and CNV of *ERBB2*. As a result, at least one tumor-derived mutation could be detected at baseline in 21/30 patients (70%), during chemotherapy in 3/32 patients (9.4%), post-chemotherapy in 9/31 patients (29.0%), and post-operatively in 5/32 patients (15.6%) (Table 1). All patients with negative ctDNA at baseline demonstrated no mutation in blood ctDNA at any other time point, indicating the importance of determining the baseline status of ctDNA (Table 1).

At baseline, premenopausal patients had significantly more positive ctDNA than menopausal patients ( $p = 0.013$ , Fig. 2A). We also evaluated the variant allele frequency (VAF) for each plasma sample to reflect the circulating tumor burden. For plasma samples with more than one mutation, the VAF was identified as the maximal one. Similarly, the premenopausal cohort showed a higher VAF than the menopausal cohort ( $p = 0.007$ , Fig. 2A). In addition, patients with more lymph involvement had a higher proportion of positive ctDNA than those with less lymph involvement ( $p = 0.011$ , Fig. 2B). Although not reaching statistical significance, a higher VAF was observed in patients with more lymph involvement ( $p = 0.073$ , Fig. 2B). Higher fractions of positive ctDNA ( $p = 0.007$ ) and VAF ( $p = 0.043$ ) were identified in patients with advanced stage than in those with early stage tumors (Fig. 2C). Furthermore, the fraction of patients with positive ctDNA associated with high recurrence risk factors, including poor differentiation, hormone receptor negative, HER2 positive, high Ki67, and metastasis (Fig. S4). The VAF was also relatively high for patients with positive progesterone receptor and metastasis (Fig. S4). Taken together, these results indicate that the baseline ctDNA status is associated with multiple clinical factors; however, they should be further validated in large-scale cohorts.

### The peri-neoadjuvant ctDNA associated with the therapeutic efficacy

The cfDNA concentration and ctDNA VAF varied at different sampling time points. At baseline, the cfDNA concentration was the lowest ( $p < 0.001$ , Fig. 2D), whereas the VAF was the highest ( $p < 0.001$ , Fig. 2D). For the dynamic change of ctDNA, four types of patients were present: (1) ctDNA was negative at baseline, during chemotherapy, and after chemotherapy ( $n = 9$ ); (2) ctDNA was positive at baseline, but negative during and after chemotherapy ( $n = 11$ , Fig. S5A); (3) ctDNA was positive at baseline and during chemotherapy ( $n = 3$ , Fig. S5B); and (4) ctDNA was positive at baseline and after chemotherapy, but was negative during chemotherapy ( $n = 7$ , Fig. S5C). For type 3 and 4, the ctDNA residue during and after chemotherapy may be a clue about the refractory disease and be associated with therapeutic efficacy.

Whereafter, we analyzed the correlation between ctDNA status and therapeutic efficacy. Miller-Payne classification was used to assess the pathological response of carcinoma in situ. For N stage (ypN) after NCT, one patient was excluded because the post-chemotherapy blood sample was unavailable. Overall, 21 patients (67.7%) exhibited negative ctDNA during and after chemotherapy, including four pCR patients (Fig. 3). The median decrease in primary tumor volume was 54.6% (range, 17.1–98.0%) and 90.2% (range, 7.7–100%) for patients with positive and negative ctDNA, respectively (Fig. 3). With respect to Miller-Payne classification, grade 4–5 was found in only 10% (1/10) of patients with positive ctDNA but over half (52.38%, 11/21) of patients with negative ctDNA ( $p = 0.046$ ; Fig. 3, Table S2). Furthermore, we found the median number of involved lymph nodes was 3.5 (range, 0–6) for patients with positive ctDNA and 1 (range, 0–9) for patients with negative ctDNA. Over 10 regional lymph nodes were involved (ypN3) in four patients with positive ctDNA. Nevertheless, no negative patients demonstrated stage ypN3 ( $p = 0.0067$ ; Fig. 3, Table S2). To determine whether baseline ctDNA status is related to mutational detection of serial ctDNA based on the aforementioned results, we next performed the same analysis in 22 patients with positive baseline ctDNA and drew similar conclusions as expected (for Miller-Payne grade,  $p = 0.0244$ ; for stage ypN,  $p = 0.0260$ ; Table S2). These results indicated that peri-neoadjuvant ctDNA status before surgery could predict the pathological decrease in carcinoma in situ, as well as the involvement of regional lymph nodes evaluated post-operatively.

## Post-operative Ctdna Status Related To Recurrence And Survival

The median follow-up was 23.2 months (range, 16–26.3 months). All patients underwent post-operative adjuvant agents, including radiotherapy ( $n = 32$ ), aromatase inhibitor ( $n = 21$ ), and trastuzumab ( $n = 9$ ). During follow-up, seven patients (21.9%) experienced distant recurrence, including five LABC and two MBC. We evaluated the prognostic discrepancy between patients grouped by different pathological factors, CEA/CA15-3 level, and serial ctDNA status. As a result, the recurrence rate and DFS

performed similarly between patients with different Luminal isoforms, Miller-Payne grade, ypN stage, pathologic response, CEA/CA15-3 level, as well as cycle 3, and post-chemotherapy ctDNA status (Fig. S6). Although not demonstrating statistical significance, patients with clinical stage  $\geq$  or positive baseline ctDNA tended to present poorer DFS compared with their counterparts (Fig. S6). Dramatically, patients with positive ctDNA after surgery had a higher risk of recurrence and worse DFS than those with negative ctDNA (Fig. 4A). All five patients with positive ctDNA post-operatively experienced distant metastasis during follow-up, while only 7.41% (2/27) of those with negative ctDNA relapsed. The median DFS was 9.8 months for the positive ctDNA cohort but not reached for the negative ctDNA cohort (HR 23.53; 95% CI, 1.904–290.9;  $p < 0.0001$ ). Considering the potential prognostic impact of MBC and baseline ctDNA status, we further performed subgroup analysis only focusing on patients with clinical stage  $\geq$  or positive baseline ctDNA. Similar to the total cohort, those with positive ctDNA after surgery also displayed poorer prognosis compared with their counterparts (for patients with clinical stage  $\geq$ , recurrence rate, 100% vs. 0%, median DFS, 9.8 months vs. not reached, HR and 95% CI were unevaluable,  $p < 0.0001$ , Fig. 4B; for patients with positive baseline ctDNA, recurrence rate, 100% vs. 11.8%, median DFS, 9.8 months vs. not reached, HR, 15.59, 95% CI 1.822–133.4,  $p < 0.0001$ , Fig. 4C).

## Sequential Ctdna Analysis In Tumor Surveillance

For seven patients with distant recurrence, blood samples were collected 3–5 times post-operatively and during follow-up. We used the VAF to depict the dynamic change of tumor burden for each patient. The CEA/CA15-3 levels at some time points were also taken into consideration. As shown in Fig. 5, the quantitative VAF of ctDNA fluctuated with the clinical performance of each patient. All patients experienced a dramatic increase in ctDNA VAF at relapse, and multiple ctDNA samples were positive between surgery and relapse, except for those of P25. P03 and P08 underwent adjuvant radiotherapy and a sequential aromatase inhibitor regimen, and ctDNA samples collected post-operatively and at recurrence were both positive. The VAF increased despite therapeutic agents and persisted in subsequent time points. Particularly for P03, a blood sample was collected at the end of chemotherapy after recurrence. The therapeutic evaluation was partial response, and a significant decrease in the VAF was observed. For P19, the VAF was stable during adjuvant radiotherapy but increased dramatically at recurrence. The therapeutic evaluation of second-line radiotherapy was PR, and the tumor progressed 112 days later, both of which were reflected in the dynamic change in ctDNA VAF. For P24 and P39, two blood samples were collected during adjuvant radiotherapy and sequential trastuzumab intake, the latter of which was detected with tumor-derived mutations. One hundred and sixty days later, both patients experienced bone or chest metastasis, and ctDNA VAF also increased to a relatively high level. For P24, the ctDNA burden then decreased significantly due to subsequent chemotherapy and lapatinib intake. For P25, only blood collected at recurrence was detectable with tumor-derived mutations. P41 underwent adjuvant radiotherapy and a sequential aromatase inhibitor regimen with trastuzumab. CtDNA VAF was initially stable and then increased when chest, lymph node, and bone metastasis arose. In consideration of the hypothesis that positive ctDNA represents the post-operative minimal residual disease before recurrence, tumor relapse could be predicted in advance with a median interval of 6.9 months (range, 0–12.1 months).

However, the sensitivity of conventional CEA/CA15-3 in tumor surveillance was unsatisfactory. Only one (P25) patients exhibited elevated CEA/CA15-3 post-operatively; two (P08, P25) had elevated CEA/CA15-3 at recurrence (Fig. 5). Furthermore, P24 and P25 showed elevated CEA/CA15-3 levels at 6.9 and 7.3 months before recurrence, respectively, indicating an advantage in relapse prediction over ctDNA (3.3 and 0 months) (Fig. 5). Altogether, the combination of CEA/CA15-3 and ctDNA may improve treatment monitoring in patients with locally advanced or metastatic breast cancer.

## Discussion

Here, we report a study about the application of ctDNA to monitor neoadjuvant chemotherapy for breast cancer. We enrolled dozens of patients with locally advanced or metastatic breast cancer and collected serial blood samples around neoadjuvant chemotherapy, after surgery, and during follow-up. In addition, tissue biopsies were collected at diagnosis and during surgery to fully reveal the patient-specific mutational landscape and traced these tumor-derived mutations in serial blood ctDNA. As a result, peri-neoadjuvant ctDNA status was related to Miller-Payne classification and ypN stage determined through post-operative pathologic evaluation, and post-operative ctDNA status was associated with survival. Furthermore, the dynamic mutational burden of ctDNA was correlated with the clinical outcomes of recurrent patients, and distant recurrence could be predicted months in

advance via the quantification of tumor-specific mutations in ctDNA. Compared with a single marker, the combination of CEA/CA15-3 and ctDNA seem to be a more promising and sensitive method in tumor surveillance.

To date, the prognosis of breast cancer has been significantly improved by multiple new agents. Based on current guidelines, neoadjuvant chemotherapy is the routine treatment for locally advanced breast cancer. This therapy has been increasingly used to shrink tumors enough to qualify for breast-conserving surgery and assess early in-vivo response to systemic therapies[21, 22]. The determination of pCR after surgery is the main surrogate marker for adjuvant agents and long-term survival[23, 24]. Nevertheless, this correlation mainly exists at an individual patient level and is not consistent at the population level. One published study failed to show a relationship between the improved pCR rate and improved DFS[25]. Some aggressive phenotypes, such as triple-negative breast cancer, display high pCR rates but also have frequently early metastatic relapses[26]. Based on aforementioned issues, two areas of research are critical to improving care for patients with breast cancer undergoing neoadjuvant chemotherapy: (1) new markers are urgently needed to predict clinical response to neoadjuvant chemotherapy to better guide treatment planning and management; and (2) in addition to pathological evaluation, how we can determine better prognostic indicator.

To solve these questions, we focused on ctDNA, a novel and tumor-specific biomarker that has been broadly explored in recent years. In breast cancer, ctDNA detection is associated with clinical stage[11]. For patients with stage I and II, the detectable rate is much higher than those with stage III and IV. In this study, most patients were stage I, accompanied by several patients at stage II and III. The detectable rate of 70% at baseline is in concordance with the previous study[11]. Patients with advanced clinical N and M stages demonstrated a higher rate than those with early stages; however, it did not reach statistical significance, perhaps due to the small sample size. In addition, we found that premenopausal patients were more likely to have positive ctDNA than older patients, which may support the notion that premenopausal breast cancer is more aggressive than post-menopausal breast cancer[26, 27]. Baseline ctDNA status was strongly associated with the detectability of serial ctDNA. In patients with negative ctDNA at baseline, no tumor-specific ctDNA detected in any blood samples. This finding suggests that the detectability of ctDNA is an inherent factor for each tumor. Therefore, determining a patient's ctDNA status at baseline is most important prior to other ctDNA-related exploration. This finding may help enhancing the sensitivity of serial ctDNA test, as well as avoiding excessive detection with higher efficiency and less medical expense.

The vast majority of patients with positive ctDNA at baseline demonstrated an obvious decline in the VAF—indicating substantial downstaging—after neoadjuvant chemotherapy. However, for some patients, tumor-derived mutations were still detectable in blood samples at cycle 3 or post-chemotherapy. When integrated with pathological evaluation, we found that these patients have a worse prognosis than those whose tumor-specific ctDNA disappeared. The significant correlation between peri-neoadjuvant ctDNA status and pathological evaluation suggests that serial ctDNA analysis around neoadjuvant treatment may be used to guide subsequent therapeutic strategies, such as escalation or de-escalation of systemic treatment. Nevertheless, the criteria of differentiation should be further explored with larger-scale studies.

Compared with baseline, ctDNA remained detectable in much fewer patients after chemotherapy (29%) and after surgery (15.6%), suggesting that ctDNA detected at baseline is mostly derived from the primary rather than the distant metastatic sites. This may explain why baseline ctDNA status is not prognostic. Namely, ctDNA can reflect minimal residual disease, which is associated with the risk of relapse, only when the primary tumor has been removed. Post-chemotherapy ctDNA is also not prognostic possibly due to the incomplete elimination of markers derived from primary sites, which can be modestly drawn from the relatively higher positive rate than post-operative ctDNA. Although pathological risk factors are the routine marker to determine adjuvant agents and prognosis, the correlation is still disputable[28]. CEA and CA15-3 are used in post-operative monitoring for breast cancer, but the sensitivity and specificity are far from satisfying[5, 6]. Of seven patients who experienced post-operative recurrence, elevated CEA/CA15-3 was observed in three patients, and ahead of image findings of tumor recurrence in two patients, including one who was undetectable with ctDNA before imaging recurrence. As a comparison, ctDNA was detected at recurrence in all seven patients, and ahead of imaging recurrence in five patients. Through a combination of these markers, oncologists can better monitor the change of tumor burden and modify clinical management during follow-up.

A previous study focused on the *TP53* mutations in blood ctDNA, and its association with efficacy and prognosis of neoadjuvant chemotherapy[13]. However, breast cancer is of high genomic heterogeneity. Even the top two mutant genes, *TP53* and *PIK3CA*, do not cover an entire cohort[29–32]. As shown in this study, 37.5% (12/32) of patients had no *TP53* or *PIK3CA* mutations in tumor

tissue. Thus, a remarkable innovation of this study is the capture panel, which includes highly-frequently mutant genes across the main solid tumors. Using such approach, oncologists would no longer need to screen the applicable patients in advance, suggesting that the population who can benefit from ctDNA profiling may be enlarged. To eliminate the interference of background and hematopoietic mutations in the blood, we profiled the mutational landscape for each patient. Using two longitudinal tissue samples, temporal heterogeneity was revealed and could be avoided to some extent in the test for ctDNA. More longitudinal and multi-regional sampling can further eliminate the interference of spatial-temporal heterogeneity, but we think our strategy benefits patients most when considering physical invasion and economic cost.

Several limitations persisted in this study, including its modest sample size and the lack of analysis about ctDNA detection in diverse patient subsets with different clinicopathological characteristics. Particularly, only four MBC were enrolled which restricted corresponding subgroup analysis. Besides, it was still confused whether post-operative ctDNA was derived from minimal residual disease or metastatic lesions, and simultaneous sampling of primary lesion, metastasis and peripheral blood might contribute to the problem. Despite aforementioned limitations, the results of this study offer a promising method to stratify patients before treatment by clinical management parameters, including surgical procedures and adjuvant agents, via noninvasive and highly sensitive blood ctDNA analysis. These results should be further explored and validated in larger-scale studies. We hope our efforts can help improve the clinical outcomes for patients undergoing neoadjuvant chemotherapy.

## Methods

### Clinical cohort

This prospective study recruited LABC patients at Peking Union Medical College Hospital (ClinicalTrials.gov Identifier: NCT02797652). All the patients were clinically evaluated by physical examination and imaging examination. Patients were enrolled when they met at least one of the following inclusion criteria: (1) tumors larger than 5 cm with regional lymphadenopathy (N1–N3); (2) cancers that involve the skin of the breast or the underlying muscles of the chest; (3) presence of regional lymphadenopathy (clinically fixed or matted axillary lymph nodes, or any infraclavicular, supraclavicular, or internal mammary lymphadenopathy) regardless of tumor stage; or (4) inflammatory breast cancer. The clinical stage was evaluated according to the American Joint Committee on Cancer TNM staging system for breast cancer[33]. In addition, some patients with metastatic breast cancer (MBC, stage  $\geq$  II) were also enrolled according to a proposal from at least three experienced clinicians. These patients had controllable metastasis, mainly in bone and internal mammary lymph nodes. For preoperative therapy regimens, all the patients with HER2- cancer were given docetaxel/epirubicin/cyclophosphamide (TEC) regimen. Although dual HER2-blockade with both trastuzumab and pertusumab was considered to be the optimal choice for HER2+ breast cancer in neoadjuvant setting, the unavailability of pertusumab in our country led to the adoption of docetaxel/carboplatin/trastuzumab (TCbH) regimen for all the HER2+ patients. And all the HER2+ patients received one year trastuzumab as adjuvant HER2-targeting therapy. The choice of surgery and medical agents was made by the treating clinician, who was blinded to the ctDNA profiling.

After surgery, the resected tissues were sent to the pathology laboratory at Peking Union Medical College Hospital. Both breast and lymph node tissue was estimated to determine ypT and ypN staging. The Miller-Payne classification system was used to evaluate the outcome of neoadjuvant chemotherapy[34]. We defined pathologic complete response (pCR) as simultaneous Miller-Payne grade 5 (no residual invasive cancer, but ductal carcinoma in situ may be present) and ypN0 (absence of cancer in lymph nodes). The concentrations of serum CEA and CA15-3 post-operatively or during follow-up were measured by the diagnostic laboratory at Peking Union Medical College Hospital. The cut-off value is 5  $\mu$ g/L for elevated CEA and 30 U/mL for elevated CA15-3.

This study was registered at ClinicalTrials.gov: Clinical Application of CTDNA in Operable Breast Cancer Patients, NCT02797652 (<https://clinicaltrials.gov/ct2/show/record/NCT02797652?term=NCT02797652&draw=2&rank=1>). Registration date: June 13, 2016. This study was approved by the ethical committee at Peking Union Medical College Hospital. Informed written consents were obtained from all participants before any study-related procedures began. This study was performed in accordance with the Declaration of Helsinki.

### Sample collection and DNA extraction

Before neoadjuvant treatment, tumor tissue biopsy was collected via core needle aspiration. In addition, we collected resected specimens during surgery. All tissue samples were fixed by formalin and then embedded in paraffin. At least 10-mL blood samples were collected in Streck tubes prior to chemotherapy, three cycles after the initiation of chemotherapy, 1–6 weeks after chemotherapy, 3–7 days after surgery, and at several follow-up time points. Less than 72 h after collection, peripheral blood was processed by centrifugation at  $1,600 \times g$  for 10 min, then transferred to microcentrifuge tubes and centrifuged again at  $16,000 \times g$  for 10 min to remove cell debris. Plasma and peripheral blood lymphocytes (PBLs) were separated and stored at  $-80^{\circ}\text{C}$  prior to DNA extraction.

Intact genomic DNA was extracted from tumor tissue and PBL using the QIAamp DNA mini kit and QIAamp DNA blood mini kit (Qiagen, Hilden, Germany), respectively. Circulating cell-free DNA (cfDNA) fragments were extracted from plasma using the QIAamp circulating nucleic acid kit (Qiagen, Hilden, Germany). Qubit fluorometer and Qubit dsDNA high sensitivity assay kit (Invitrogen, Carlsbad, CA, USA) were used to estimate the concentration of DNA. CfDNA fragment length was also estimated using an Agilent 2100 bioanalyzer and the DNA high sensitivity kit (Agilent Technologies, Santa Clara, CA, USA) to evaluate the contamination of large DNA fragments.

## Sequencing library construction

Before library construction, genomic DNA was sheared into 200–250-bp fragments using the Covaris S2 instrument (Woburn, MA, USA). The DNA library preparation kit for Illumina (New England Biolabs, Ipswich, MA, USA) was used to prepare indexed next-generation sequencing libraries for genomic DNA fragments. As for cfDNA, after end-repairing and A-tailing reactions, well-designed adapters with unique identifiers were ligated to both ends of double-stranded cfDNA fragments. Then, PCR was used to generate sufficient fragments prior to hybridization. A previous article described additional detailed information about library construction[35].

## Target region capture and next-generation sequencing

In hybrid capture procedure, we recruited custom-designed biotinylated oligonucleotide probes (IDT, Coralville, IA, USA) covering 1.09 Mbp of the human genome. Several processes were performed to embody the capture panel: (1) the most common driver genes across 12 common solid tumors[36] were selected and their entire exome regions were chosen; (2) genomic regions related with efficacy of chemotherapy, targeted drugs, and immunotherapy were added according to available clinical and pre-clinical studies; and (3) high-frequently mutant regions recorded in the Catalogue of Somatic Mutations in Cancer (COSMIC, <http://cancer.sanger.ac.uk/cosmic>) and The Cancer Genome Atlas (TCGA, <https://cancergenome.nih.gov/>) were added. The overall included genes and genomic regions were shown in *Table S1*. We performed DNA sequencing using an Illumina 2 × 75-bp paired-end sequencing strategy on the HiSeq Sequencing System (Illumina, San Diego, CA, USA). The average sequencing depth was 800× for tissue DNA, 200× for PBL DNA, and 1,500× for cfDNA. Additional experimental information about target region capture and sequencing was also described in a previous article[35].

## Raw data processing

Raw reads with adaptor sequences were first removed. Afterward, reads with more than 50% low-quality bases or more than 50% N bases together with their mate pair were further filtered. Residual reads were mapped to the reference human genome (hg19) using the Burrows-Wheel Aligner (<http://bio-bwa.sourceforge.net/>) with default parameters. Duplicate reads were identified and marked with Picard's Mark Duplicates tool ([https://software.broadinstitute.org/gatk/documentation/tooldocs/4.0.3.0/picard\\_sam\\_markduplicates\\_MarkDuplicates.php](https://software.broadinstitute.org/gatk/documentation/tooldocs/4.0.3.0/picard_sam_markduplicates_MarkDuplicates.php)). CfDNA data were clustered according to unique identifiers and the position of the template fragments. False-positive events introduced by PCR or sequencing procedures were removed according to clustered reads and mate pair reads. Local realignment and base quality recalibration were performed using the Gene Analysis Toolkit (<https://www.broadinstitute.org/gatk/>).

# Somatic mutation calling

Somatic single-nucleotide variations (SNVs) were called using the MuTect2 algorithm

([https://software.broadinstitute.org/gatk/documentation/tooldocs/3.8-](https://software.broadinstitute.org/gatk/documentation/tooldocs/3.8-0/org_broadinstitute_gatk_tools_walkers_cancer_m2_MuTect2.php)

[0/org\\_broadinstitute\\_gatk\\_tools\\_walkers\\_cancer\\_m2\\_MuTect2.php](https://software.broadinstitute.org/gatk/documentation/tooldocs/3.8-0/org_broadinstitute_gatk_tools_walkers_cancer_m2_MuTect2.php)). Candidate mutations were removed according to several criteria: (1) more than 10 reads with insertions/deletions in an 11-bp window were centered; (2) the matched DNA derived from PBL control sample carried  $\geq 3\%$  or  $\geq 2\%$  alternate allele reads, and the sum of quality scores was above 80; (3) the candidate was found in several single-nucleotide polymorphism databases (dbSNP, <https://www.ncbi.nlm.nih.gov/projects/SNP/>; 1000G, <https://www.1000genomes.org/>; ESP6500, <https://evs.gs.washington.edu/>; ExAC, <http://exac.broadinstitute.org/>), but not listed in the COSMIC database; (4) the candidate was supported by fewer than five high-quality reads (base quality  $\geq 30$ , mapping quality  $\geq 30$ ); and (5) the allele frequency was less than 1% for genomic DNA.

Insertions or deletions of small fragments (Indels) were called using MuTect2 with default parameters. Variants were removed if they were detected in matched control samples with three or more reads indicating Indels at the same location or in 40-bp flanking regions of experimental samples or residing near regions with low complexity or short tandem repeats.

Copy number variants (CNVs) of *ERBB2* were detected using the Contra algorithm (<http://contra-cnv.sourceforge.net>).

## Statistical analysis

Descriptive statistics were performed using SPSS 22.0 (IBM, Armonk, NY, USA). Differences in baseline characteristics between patients with positive and negative ctDNA at multiple time points were estimated using Fisher's exact test for categorical variables and Mann-Whitney (rank sum) test for continuous variables (SPSS 22.0, IBM, Armonk, NY, USA). Differences of cfDNA concentration and ctDNA allele frequency collected at different time points were calculated using One-Way ANOVA analysis. The primary endpoint was disease-free survival, measured from surgery to the first relapse or until the most recent follow-up of this study. A Kaplan-Meier analysis with the log-rank test was used to compare the disease-free survival (DFS) of different subsets of patients (GraphPad Prism 7, GraphPad Software, La Jolla, CA, USA). A *p*-value less than 0.05 indicated statistical significance.

## Declarations

### Data Availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

### Author Contributions

Totally 14 authors are listed in manuscript. Yidong Zhou and Yaping Xu are the main experimental designers and writers; Changjun Wang, Ru Yao and Li Peng have assisted the management of patients and collection of samples; Yuhua Gong, Yanyan Zhang, Xiuli Zhu and Jing Bai helped data arrangement and analysis; Yanfang Guan, Xuefeng Xia, Ling Yang and Xin Yi offered constructive suggestion for study design and analytical methods.

### Competing Interests statement

The authors declare that there are no competing interests.

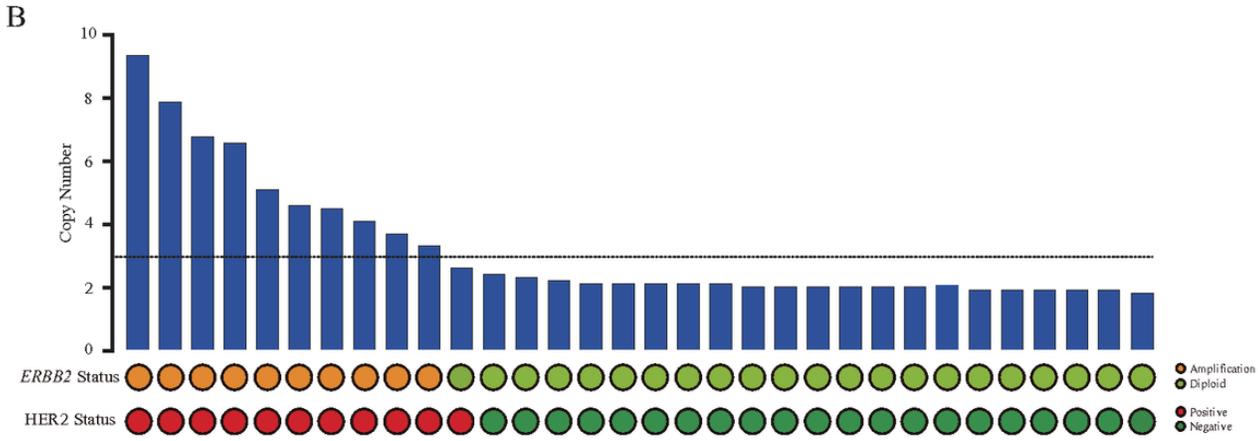
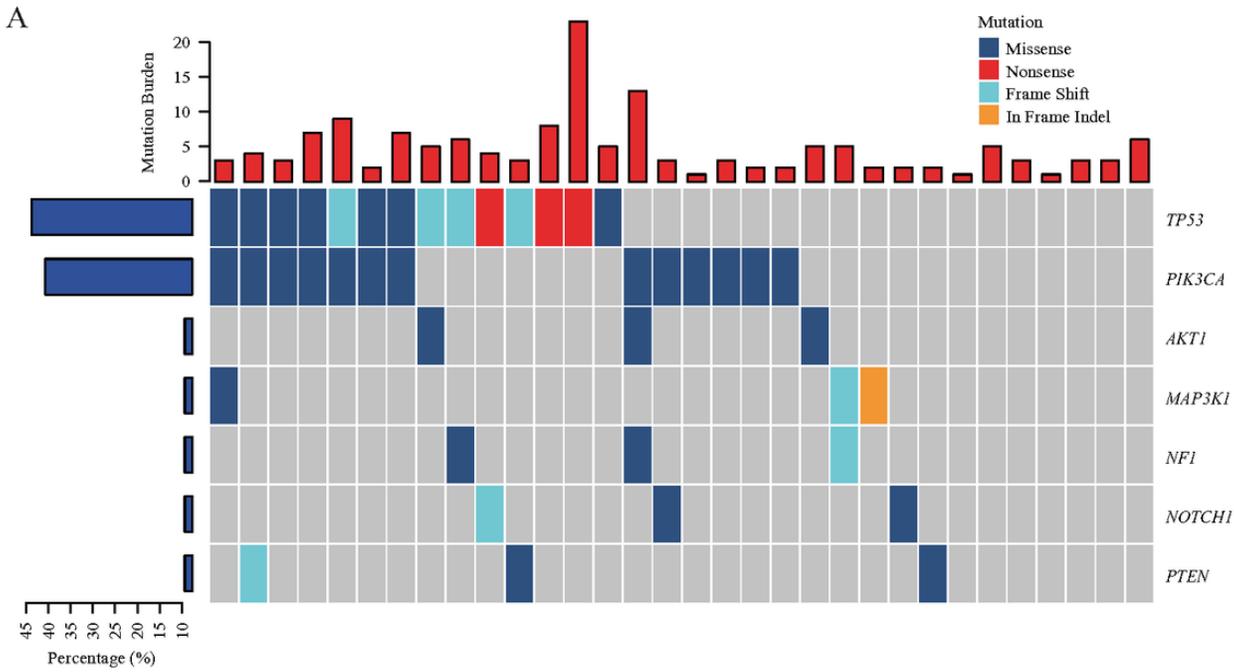
## References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A (2018) Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *Cancer J Clin* 68(6):394–424
2. King TA, Morrow M (2015) Surgical issues in patients with breast cancer receiving neoadjuvant chemotherapy. *Nature reviews Clinical oncology* 12(6):335–343

3. Alawad AA (2014) Evaluation of clinical and pathological response after two cycles of neoadjuvant chemotherapy on Sudanese patients with locally advanced breast cancer. *Ethiopian journal of health sciences* 24(1):15–20
4. Takahashi H, Kagara N, Tanei T, Naoi Y, Shimoda M, Shimomura A, Shimazu K, Kim SJ, Noguchi S (2017) Correlation of Methylated Circulating Tumor DNA With Response to Neoadjuvant Chemotherapy in Breast Cancer Patients. *Clin Breast Cancer* 17(1):61–69 e63
5. Bottini A, Berruti A, Tampellini M, Morrica B, Brunelli A, Gnocchi E, Brizzi MP, Aguggini S, Fara E, Alquati P et al (1997) Influence of neoadjuvant chemotherapy on serum tumor markers CA 15 – 3, MCA, CEA, TPS and TPA in breast cancer patients with operable disease. *Tumour Biol* 18(5):301–310
6. Wang YJ, Huang XY, Mo M, Li JW, Jia XQ, Shao ZM, Shen ZZ, Wu J, Liu GY (2015) Serum Tumor Marker Levels might have Little Significance in Evaluating Neoadjuvant Treatment Response in Locally Advanced Breast Cancer. *Asian Pacific journal of cancer prevention: APJCP* 16(11):4603–4608
7. Zhou X, Li Y (2016) Local Recurrence after Breast-Conserving Surgery and Mastectomy Following Neoadjuvant Chemotherapy for Locally Advanced Breast Cancer - a Meta-Analysis. *Breast care* 11(5):345–351
8. Vargo JA, Beriwal S, Ahrendt GM, Soran A, Johnson RR, McGuire K, Bhargava R (2011) Molecular class as a predictor of locoregional and distant recurrence in the neoadjuvant setting for breast cancer. *Oncology* 80(5–6):341–349
9. Wan JCM, Massie C, Garcia-Corbacho J, Moulriere F, Brenton JD, Caldas C, Pacey S, Baird R, Rosenfeld N (2017) Liquid biopsies come of age: towards implementation of circulating tumour DNA. *Nat Rev Cancer* 17(4):223–238
10. Siravegna G, Marsoni S, Siena S, Bardelli A (2017) Integrating liquid biopsies into the management of cancer. *Nature reviews Clinical oncology* 14(9):531–548
11. Bettegowda C, Sausen M, Leary RJ, Kinde I, Wang Y, Agrawal N, Bartlett BR, Wang H, Luber B, Alani RM et al (2014) Detection of circulating tumor DNA in early- and late-stage human malignancies. *Sci Transl Med* 6(224):224ra224
12. Garcia-Murillas I, Schiavon G, Weigelt B, Ng C, Hrebien S, Cutts RJ, Cheang M, Osin P, Nerurkar A, Kozarewa I et al (2015) Mutation tracking in circulating tumor DNA predicts relapse in early breast cancer. *Sci Transl Med* 7(302):302ra133
13. Riva F, Bidard FC, Houy A, Saliou A, Madic J, Rampanou A, Hego C, Milder M, Cottu P, Sablin MP et al (2017) Patient-Specific Circulating Tumor DNA Detection during Neoadjuvant Chemotherapy in Triple-Negative Breast Cancer. *Clinical chemistry* 63(3):691–699
14. Kim JY, Park D, Son DS, Nam SJ, Kim SW, Jung HH, Kim YJ, Park G, Park WY, Lee JE et al (2017) Circulating tumor DNA shows variable clonal response of breast cancer during neoadjuvant chemotherapy. *Oncotarget* 8(49):86423–86434
15. Chen YH, Hancock BA, Solzak JP, Brinza D, Scafe C, Miller KD, Radovich M (2017) Next-generation sequencing of circulating tumor DNA to predict recurrence in triple-negative breast cancer patients with residual disease after neoadjuvant chemotherapy. *NPJ breast cancer* 3:24
16. Lehner J, Stotzer OJ, Fersching D, Nagel D, Holdenrieder S (2013) Circulating plasma DNA and DNA integrity in breast cancer patients undergoing neoadjuvant chemotherapy. *Clin Chim Acta* 425:206–211
17. Zardavas D, Irrthum A, Swanton C, Piccart M (2015) Clinical management of breast cancer heterogeneity. *Nature reviews Clinical oncology* 12(7):381–394
18. Janiszewska M, Liu L, Almendro V, Kuang Y, Paweletz C, Sakr RA, Weigelt B, Hanker AB, Chandralapaty S, King TA et al (2015) In situ single-cell analysis identifies heterogeneity for PIK3CA mutation and HER2 amplification in HER2-positive breast cancer. *Nat Genet* 47(10):1212–1219
19. Bruna A, Rueda OM, Greenwood W, Batra AS, Callari M, Batra RN, Pogrebniak K, Sandoval J, Cassidy JW, Tufegdzcic-Vidakovic A et al (2016) A Biobank of Breast Cancer Explants with Preserved Intra-tumor Heterogeneity to Screen Anticancer Compounds. *Cell* 167(1):260–274 e222
20. De Mattos-Arruda L, Cortes J, Santarpia L, Vivancos A, Tabernero J, Reis-Filho JS, Seoane J (2013) Circulating tumour cells and cell-free DNA as tools for managing breast cancer. *Nature reviews Clinical oncology* 10(7):377–389
21. van Nes JG, Putter H, Julien JP, Tubiana-Hulin M, van de Vijver M, Bogaerts J, de Vos M, van de Velde CJ (2009) Cooperating Investigators of the E: Preoperative chemotherapy is safe in early breast cancer, even after 10 years of follow-up; clinical and translational results from the EORTC trial 10902. *Breast cancer research treatment* 115(1):101–113

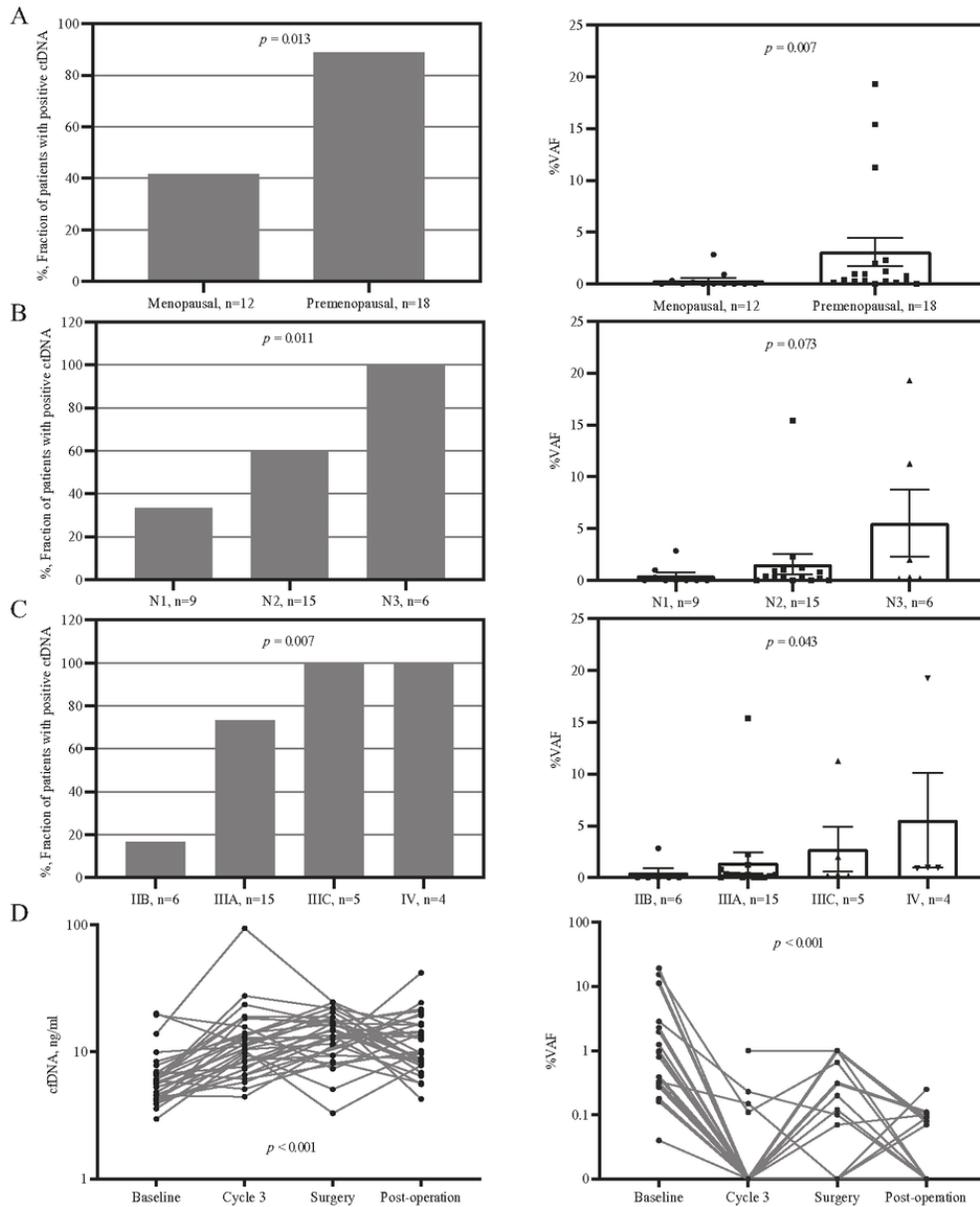
22. von Minckwitz G, Untch M, Blohmer JU, Costa SD, Eidtmann H, Fasching PA, Gerber B, Eiermann W, Hilfrich J, Huober J et al (2012) Definition and impact of pathologic complete response on prognosis after neoadjuvant chemotherapy in various intrinsic breast cancer subtypes. *Journal of clinical oncology: official journal of the American Society of Clinical Oncology* 30(15):1796–1804
23. von Minckwitz G, Untch M, Nuesch E, Loibl S, Kaufmann M, Kummel S, Fasching PA, Eiermann W, Blohmer JU, Costa SD et al (2011) Impact of treatment characteristics on response of different breast cancer phenotypes: pooled analysis of the German neo-adjuvant chemotherapy trials. *Breast Cancer Res Treat* 125(1):145–156
24. Cortazar P, Zhang L, Untch M, Mehta K, Costantino JP, Wolmark N, Bonnefoi H, Cameron D, Gianni L, Valagussa P et al (2014) Pathological complete response and long-term clinical benefit in breast cancer: the CTNeoBC pooled analysis. *Lancet* 384(9938):164–172
25. Denkert C, von Minckwitz G, Darb-Esfahani S, Lederer B, Heppner BI, Weber KE, Budczies J, Huober J, Klauschen F, Furlanetto J et al (2018) Tumour-infiltrating lymphocytes and prognosis in different subtypes of breast cancer: a pooled analysis of 3771 patients treated with neoadjuvant therapy. *Lancet Oncol* 19(1):40–50
26. Rosenberg SM, Partridge AH (2015) Management of breast cancer in very young women. *Breast* 24(Suppl 2):S154–S158
27. Wang-Lopez Q, Chalabi N, Abrial C, Radosevic-Robin N, Durando X, Mouret-Reynier MA, Benmammar KE, Kullab S, Bahadour M, Chollet P et al (2015) Can pathologic complete response (pCR) be used as a surrogate marker of survival after neoadjuvant therapy for breast cancer? *Crit Rev Oncol Hematol* 95(1):88–104
28. Luen S, Virassamy B, Savas P, Salgado R, Loi S (2016) The genomic landscape of breast cancer and its interaction with host immunity. *Breast* 29:241–250
29. Pereira B, Chin SF, Rueda OM, Vollan HK, Provenzano E, Bardwell HA, Pugh M, Jones L, Russell R, Sammut SJ et al (2016) The somatic mutation profiles of 2,433 breast cancers refines their genomic and transcriptomic landscapes. *Nature communications* 7:11479
30. Cancer Genome Atlas N (2012) Comprehensive molecular portraits of human breast tumours. *Nature* 490(7418):61–70
31. Ciriello G, Gatza ML, Beck AH, Wilkerson MD, Rhie SK, Pastore A, Zhang H, McLellan M, Yau C, Kandoth C et al (2015) Comprehensive Molecular Portraits of Invasive Lobular Breast Cancer. *Cell* 163(2):506–519
32. Zehir A, Benayed R, Shah RH, Syed A, Middha S, Kim HR, Srinivasan P, Gao J, Chakravarty D, Devlin SM et al (2017) Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. *Nature medicine* 23(6):703–713
33. Singletary SE, Allred C, Ashley P, Bassett LW, Berry D, Bland KI, Borgen PI, Clark G, Edge SB, Hayes DF et al (2002) Revision of the American Joint Committee on Cancer staging system for breast cancer. *Journal of clinical oncology: official journal of the American Society of Clinical Oncology* 20(17):3628–3636
34. Ogston KN, Miller ID, Payne S, Hutcheon AW, Sarkar TK, Smith I, Schofield A, Heys SD (2003) A new histological grading system to assess response of breast cancers to primary chemotherapy: prognostic significance and survival. *Breast* 12(5):320–327
35. Lv X, Zhao M, Yi Y, Zhang L, Guan Y, Liu T, Yang L, Chen R, Ma J, Yi X: Detection of Rare Mutations in CtDNA Using Next Generation Sequencing. *Journal of visualized experiments: JoVE* 2017(126)
36. Kandoth C, McLellan MD, Vandin F, Ye K, Niu B, Lu C, Xie M, Zhang Q, McMichael JF, Wyczalkowski MA et al (2013) Mutational landscape and significance across 12 major cancer types. *Nature* 502(7471):333–339

## Figures



**Figure 1**

Mutational prevalence of tumor tissue. (A) Top seven genes mutated in over three tissue specimens. The union set of baseline tissue obtained via needle aspiration and surgical tissue is shown. The upper bars indicate the number of mutations for each patient. The left bars represent the frequency of each mutant gene. (B) Overview of clinical HER2 status and ERBB2 amplification for 32 included patients. The upper bars illustrate the ERBB2 copy number. A copy number greater than 3 was defined as amplification, which is indicated by the dotted line. Indel; insertions or deletions of small fragments.



**Figure 2**

The change of ctDNA parameters according to baseline clinical characteristics and sampling nodes. In clinical characteristic-related analysis, patients were grouped according to (A) differentiation, (B) individual N stage and (C) clinical stage. For the comparison of ctDNA detectability, the p-value was calculated by Fisher's exact test; for the comparison of ctDNA VAF, the p-value was calculated by Mann-Whitney test. (D) Change of cfDNA concentration and ctDNA VAF at different time points. The p-value was calculated by one-way ANOVA.  $P < 0.05$  indicated statistical significance.

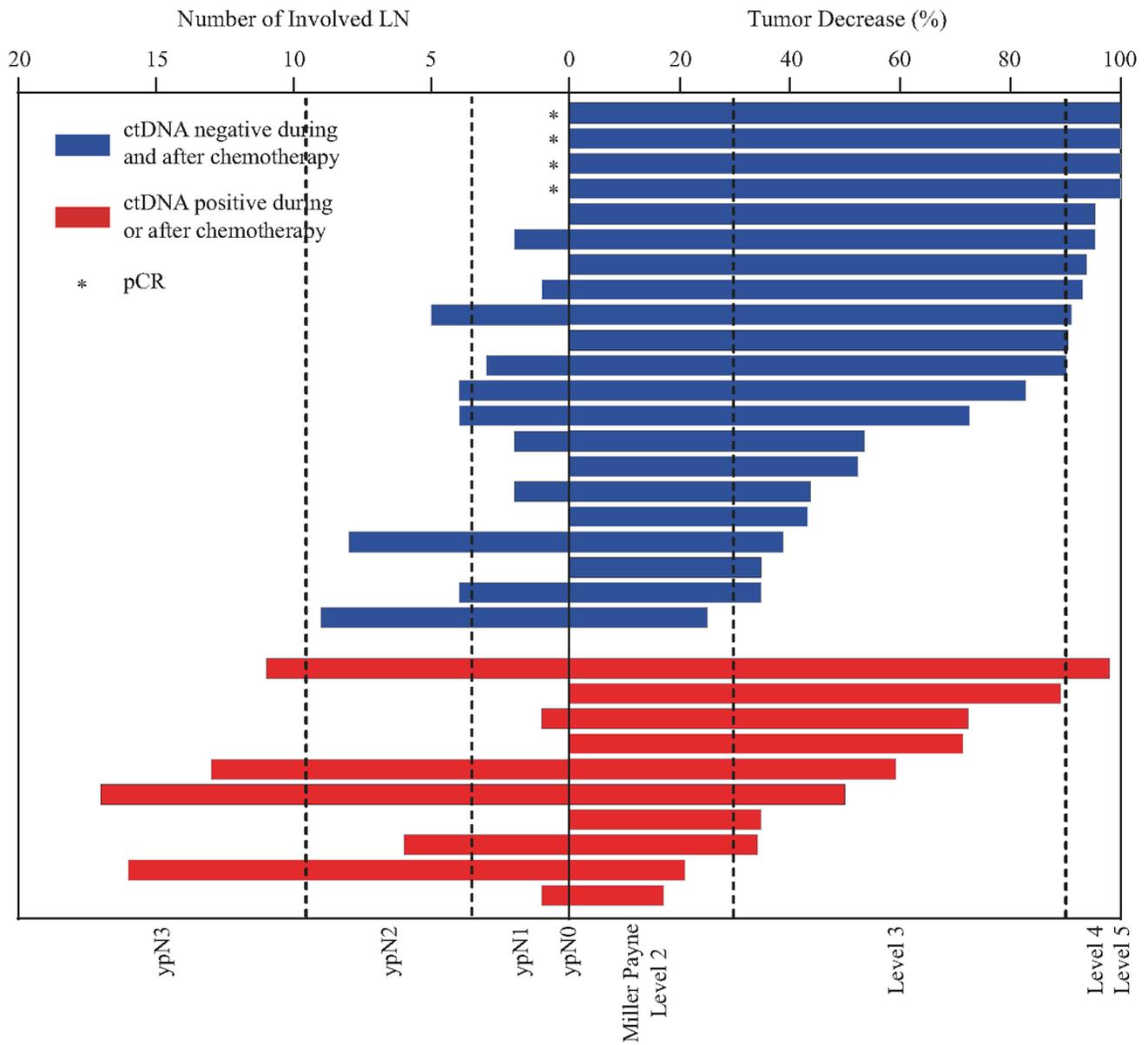


Figure 3

Correlation between ctDNA status and pathological evaluation. Pathologic complete response is defined as Miller-Payne grade 5 and ypN0.

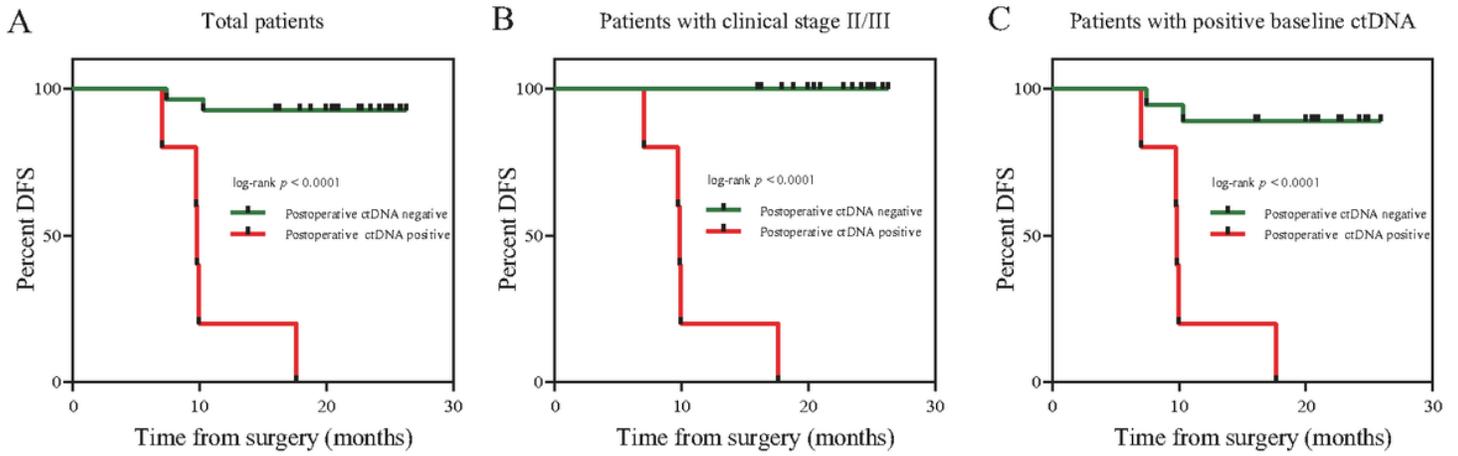


Figure 4

Prognostic survival according to post-operative ctDNA status. The whole cohort (A), patients with clinical stage II/III (B), and patients with positive ctDNA at baseline (C). NCT, neoadjuvant chemotherapy. The p-values were calculated by Kaplan-Meier survival analysis.  $P < 0.05$  indicated statistical significance.

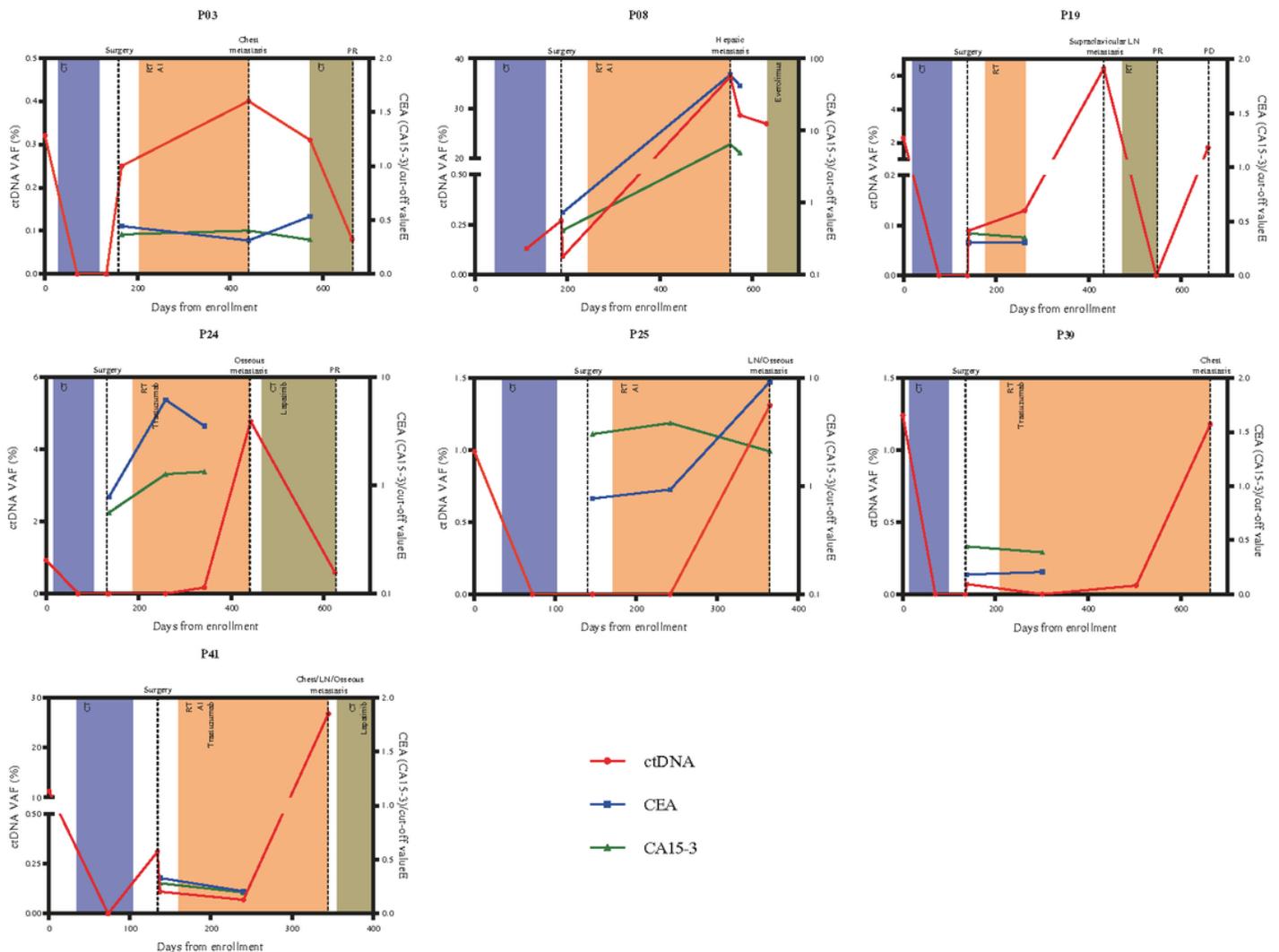


Figure 5

Dynamic change in ctDNA VAF, CEA, and CA15-3 during the entire clinical management. The level of CEA and CA15-3 were normalized according to the value of 5 µg/L and 30 U/mL, respectively.