BMC Medicine

RESEARCH

Open Access

Exploration of the clonal evolution and construction of the tumor clonal evolution rate as a prognostic indicator in metastatic breast cancer

Dan Lv[1†], Bo Lan[1†], Qihan Guo[2†], Zongbi Yi[1], Haili Qian[3], Yanfang Guan[4], Xuenan Peng[1], Ting Chen[2*] and Fei Ma[1*]

Abstract

Background Tumor heterogeneity and clonal evolution are related to the treatment resistance and disease progression in metastatic breast cancer (MBC). However, the process of clonal evolution and their relationship to prognosis remain unclear. This study aimed to elucidate the evolution of MBC through circulating tumor DNA (ctDNA) analysis and to develop a novel indicator for predicting treatment efficacy and prognosis.

Methods This multicenter retrospective study enrolled MBC patients who underwent next-generation sequencing between April 2016 and October 2022. The clonal evolution of tumors was inferred using PyClone and CITUP software.

Results The study included 406 MBC patients. A cohort of 139 patients from the National Cancer Center served as the training cohort, while 267 patients from other centers comprised the validation cohort. In the training cohort, clonal analysis revealed that most MBCs exhibited branched clonal evolution, while a minority showed linear evolution. The branched evolution pattern was associated with slower disease progression (HR, 0.53; 95% CI, 0.32–0.87; $P=0.012$). We introduced tumor clonal evolution rate (TER) as a novel concept to reflect the speed of clonal evolution. Survival analysis demonstrated that compared to the TER-high group, patients in the TER-low group had better progression-free survival (PFS) (HR, 0.62; 95% CI, 0.40–0.96; $P=0.033$) and overall survival (OS) (HR, 0.45; 95% CI, 0.24–0.85; $P=0.013$). Similarly, in the validation cohort, although the median OS was not reached, patients in the TER-low group had better prognosis compared to those in the TER-high group (HR, 0.41; 95% CI, 0.21–0.83; $P<0.001$).

Conclusions Patients with branched evolution have better treatment efficacy than those with linear evolution. The TER shows potential as a biomarker for treatment efficacy and prognosis, providing new evidence that ctDNA is a valuable molecular indicator for predicting treatment outcomes in metastatic breast cancer.

Keywords Breast cancer, Circulating tumor DNA, Clonal evolution, Tumor evolution rate

†Dan Lv, Bo Lan, and Qihan Guo contributed equally to this work.

*Correspondence:
Ting Chen
tingchen@tsinghua.edu.cn
Fei Ma
drmafei@126.com
Full list of author information is available at the end of the article

Background

Breast cancer is the most common malignant tumor affecting women's health, with approximately 2.3 million new cases worldwide in 2020 [1]. Recurrence and metastasis may occur in 20–30% of breast cancer patients after initial diagnosis or adjuvant therapy and account for approximately 90% of breast cancer-related deaths [2]. The 5-year survival rate for patients with metastatic breast cancer is only approximately 25% [3]. The dynamic evolution of tumor biology and the heterogeneity within and among tumors pose significant challenges in breast cancer treatment. Therefore, assessing tumor heterogeneity holds crucial clinical value for diagnosis, prognosis evaluation, and treatment monitoring.

Previous studies suggest that a single diagnostic tissue biopsy may not provide an accurate and comprehensive representation of the dynamic evolutionary characteristics of tumors. Multifocal and repeat biopsies can significantly enhance our understanding of tumor heterogeneity. However, the increased risk of complications and invasiveness associated with these procedures limits their widespread clinical application [4]. In the rapidly advancing field of genomics, circulating tumor DNA (ctDNA) has emerged as a blood-based biomarker analysis method, overcoming the limitations of traditional histological biopsies. ctDNA has the advantages of being less invasive and allowing for real-time and dynamic monitoring of tumor heterogeneity [5, 6]. Previous research suggests that ctDNA can detect tumor progression earlier than imaging examination [7] and cancer antigen 15–3 (CA15-3)[8], making it valuable for early tumor diagnosis. Additionally, it can serve as an indicator to assess the effectiveness of anti-tumor treatments in patients with recurrent and metastatic cancer [9–11]. Examining the mutation spectrum and frequency in ctDNA provides insight into the origin, evolutionary trajectory, and clonal composition of tumor cells. As ctDNA carries information about tumor cell mutations, this analysis provides valuable insights into the complex characteristics of tumor cells [6].

Current ctDNA research in breast cancer focuses on detection of early recurrence, identification of minimal residual disease (MRD), and evaluation of treatment response [12–15]. However, a comprehensive understanding of clonal structural changes during the progression of metastatic breast cancer is still incomplete. Therefore, second-generation sequencing technology was used in this study to identify ctDNA in various peripheral blood samples from patients. Our objective was to unveil the evolutionary process of tumors through a comparative analysis of genomic changes before and after treatment in individuals with metastatic breast cancer. Additionally, based on the changes in clonal evolution rate during the progression of breast cancer, we propose a novel indicator for determining the rate of clonal tumor evolution and explore its value in predicting antitumor treatment efficacy and prognosis.

Methods

Patients collection

We designed this retrospective, multicenter study to explore the clonal evolution of metastatic breast cancer through ctDNA analysis. The primary cohort included patients diagnosed with metastatic breast cancer who received treatment at the National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences, and Peking Union Medical College between April 25, 2016, and May 31, 2021, and who voluntarily participated in biomarker research. Patients in the validation cohort were recruited from Geneplus Medical Laboratory (Beijing, China) between April 25, 2016, and October 31, 2022.

The study included patients who met the following criteria: [1] diagnosis of distant metastatic breast cancer confirmed by cytological or histological examination at the aforementioned center; (2) peripheral blood specimens sufficient for at least two ctDNA tests; (3) complete clinical and pathological data, (4) voluntary informed consent for participation in the study. The exclusion criteria were as follows: (1) incomplete clinical or pathological data, (2) patients with locally advanced breast cancer undergoing neoadjuvant therapy, and (3) a lack of recurrence or metastasis after breast cancer surgery. The clinical and pathological data of each patient, along with relevant information about the anticancer treatments given after ctDNA testing, were meticulously documented. Ethical approval was obtained from the Ethics Review Committee of the Cancer Hospital of the Chinese Academy of Medical Sciences before commencing the study. In addition, all participating patients provided voluntary consent by signing informed consent forms. The patient enrollment flowchart is shown in Fig. 1A.

To evaluate the response to treatment, computed tomography (CT) scans were performed after every two treatment cycles or when signs or symptoms indicating disease progression were observed, following the Response Evaluation Criteria in Solid Tumors (RECIST) 1.1 guidelines.

Samples collection and plasma ctDNA testing

Peripheral blood (10 mL) was collected from each patient into Streck tubes (Streck, Omaha, NE, USA). The samples were centrifuged within 72 h to separate plasma from blood cells. Circulating DNA was then extracted from the plasma using the QIAamp Circulating Nucleic Acid

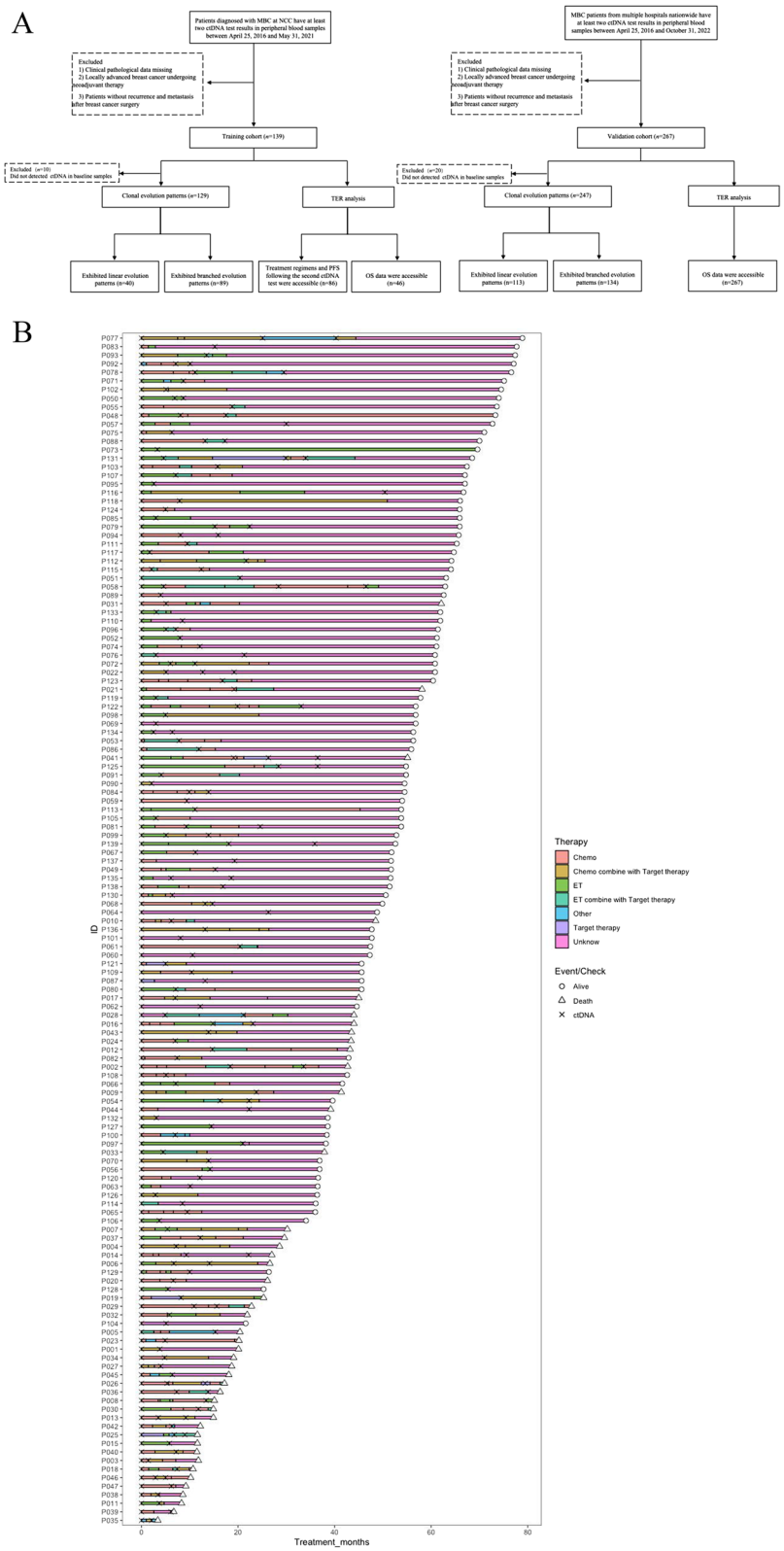

Fig. 1 Patients enrollment flowchart and treatment timeline diagram. **A** Patients enrollment flowchart. **B** ctDNA testing and treatment swimmer plots

Kit (Qiagen) and further purified with the DNeasy Blood and Tissue Kit. The size distribution of the cell-free DNA (cfDNA) was analyzed using an Agilent 2100 Bioanalyzer with the DNA HS Kit (Agilent Technologies, Santa Clara, CA, USA). Samples with a total DNA amount greater than 1 μg and without obvious degradation were considered qualified and were subjected to 1021 gene sequencing [16, 17] (Additional file 1).

Visualization of clonal evolution

PyClone was used to infer the clonal composition of patients[18]. For each patient, all genome sequencing samples collected at different time points were jointly used as input for PyClone. We used PyClone (v.0.13.1) with a beta-binomial emission density. PyClone was run with 10,000 iterations and a burn-in of 1000. The following default settings were applied for the remaining parameters: base measure parameters: alpha = 1 and beta = 1; concentration parameters: initial value = 1, prior shape = 1, and prior rate = 0.001; and beta-binomial precision parameters: initial value = 1000, prior shape = 1, prior rate = 0.001, and proposal precision = 0.01.

To reconstruct the clonal phylogenetic trees of patients, we utilized CITUP[19]. The output from PyClone, which includes the clonal assignment and cancer cell fraction (CCF) for each mutation, served as the input for CITUP. We employed the QIP version of CITUP with the following default parameters: min nodes = 1, max nodes = 8. Then, we employed the R package Timescape (https://github.com/shahcompbio/timescape) to visualize the output of CITUP, which represents the clonal evolution tree of the patient.

Definition of the TER

In tumor heterogeneity analysis, allele frequencies are commonly used to reflect the clonal composition of tumors. The variant allele frequency (VAF) of a mutation is defined as the number of mutant molecules divided by the total number of molecules containing the corresponding allele. Mutations with higher allele frequencies typically represent major clones, while those with low allele frequencies are likely to represent subclones. Our inference of the tumor phylogenetic tree for patients revealed that major clones tend to have a dominant position in the changing landscape under treatment and natural selection. To reflect the clonal evolution speed of tumors over time, considering the distribution between major and subclones, we innovatively introduced the tumor clonal evolution rate (TER) of a patient, defined as $TER = (AFmax_2/U_2 - AFmax_1/U_1)/t$.

U is the arithmetic mean of allele frequencies (AF) for all somatic mutations, while AFmax is the maximum allele frequency for somatic mutations. $AFmax_1$ and U_1 are indicators at the first detection time point T1, while $AFmax_2$ and U_2 are indicators at the second detection time point T2. t is the time interval between the two detection points, which is T2 − T1, measured in days. In our analysis, the patient's first and second ctDNA detection time points are T1 and T2, respectively.

AFmax/U is a normalized indicator whose value is not affected by tumor cell purity. When a patient has only one clone (i.e., all somatic mutations have similar allele frequencies), AFmax/U will be close to the minimum value of 1. Conversely, when a patient has multiple clones (i.e., there are significant differences in allele frequencies among somatic mutations), AFmax/U will be significantly greater than 1, and the value will be larger with the presence of more subclones (i.e., more mutations with relatively low allele frequencies). Therefore, the value of AFmax/U can reflect the tumor heterogeneity of the patient. TER characterizes the development speed of the patient's tumor heterogeneity.

Statistical analysis

Progression-free survival (PFS) was defined as the duration from the initiation of a new anticancer treatment regimen following the last ctDNA test to the date of disease progression or death. Overall survival (OS) was defined as the duration from the initiation of a new anticancer treatment regimen following the last ctDNA test to the date of death from any cause. Data for patients who were not followed up until the endpoint event (disease progression or death) were censored, and the date of the last follow-up was recorded. The latest follow-up was conducted in August 2022. All the statistical analyses were conducted using SPSS 22.0 (IBM Corp., New York, USA), GraphPad Prism 8.0 (GraphPad Software, La Jolla, CA, USA), R (v4.0.4), and R Studio (v1.4.1717). Descriptive statistics were utilized to summarize the clinical and pathological characteristics of the patients. The chi-square test or Fisher's exact test was used when comparing categorical variables. The survminer R package was used to identify the optimal threshold for TER according to OS. Survival curves were plotted using the Kaplan–Meier method, and differences in PFS and OS between treatment groups were compared using the log-rank test. Cox regression analysis was performed to compare the relationships between various clinical characteristics and PFS and OS. $P < 0.05$ was considered to indicate statistical significance.

Results

Patient characteristics and sample information

This study training cohort included 139 female patients with metastatic breast cancer. The median age was 45 years, ranging from 27 to 68 years. Among these

patients, 92.1% (128/139) had invasive ductal carcinoma, 4.3% (6/139) had invasive lobular carcinoma, and 3.6% (5/139) had other pathological types. There were 108 (77.7%) hormone receptor (HR)-positive patients and 31 (22.3%) HR-negative patients. Additionally, 28.1% (39/139) of patients were human epidermal growth factor receptor 2 (HER2)-positive, while 71.9% (100/139) were HER2-negative. Further baseline characteristics of the patients were detailed in Table 1.

ctDNA detection and dynamic mutation spectrum

Among the 139 patients, 103 underwent two rounds of ctDNA testing, 32 underwent three rounds, and 4 underwent four rounds. ctDNA testing and treatment swimmer plots were shown in Fig. 1B. ctDNA were detected in the baseline samples of 129 (92.8%) out of 139 patients. Genomic data were collected from ctDNA tests at two time points: baseline (T1) and after disease progression following treatment regimens (T2). At T1, ctDNA mutations were detected in 109 patients (84.5%) (Additional file 2: Fig. S1A). The most frequently mutated genes at T1 were TP53 (42.6%), PIK3CA (41.9%), ERBB2 (17.1%), ESR1 (14.0%), and MLL3 (13.2%), consistent with previous studies. We observed that 76.9% of triple-negative breast cancer (TNBC) harbor TP53 mutations, while the mutation frequencies in HER2-positive patients and HR-positive patients were 44.4% and 36.3%, respectively. The frequencies of PIK3CA mutations were 42.5% in HR-positive patients, 41.7% in HER2-positive patients, and 38.5% in TNBC patients, respectively. All 18 patients (18.2%) with ESR1 mutations were found among HR-positive patients who had received prior endocrine therapy before baseline ctDNA testing. At T2, the frequency of ctDNA mutations increased to 89.9%. The top mutated genes at T2 were TP53 (48.8%), PIK3CA (44.2%), ERBB2 (18.6%), ESR1 (14.7%), and GATA3 (14.7%) (Additional file 2: Fig. S1B). Comparison of mutation frequencies at T1 and T2 revealed that while the top four genes remained the same, their mutation frequencies increased at T2 (Additional file 2: Fig. S2).

Table 1 The baseline characteristics of patients

Characteristics	No. ($n=139$)[a]	Percentage (%)
Age at diagnosis		
≤ 35	14	10.1
35–60	115	82.7
> 60	10	7.2
Pathological type		
Invasive ductal carcinoma	128	92.1
Invasive lobular carcinoma	6	4.3
Others	5	3.6
HR status		
Positive	108	77.7
Negative	31	22.3
HER2 status		
Positive	39	28.1
Negative	100	71.9
Molecular subtype		
HR + /HER2 –	87	62.6
HER2 +	39	28.0
HR – /HER2 –	13	9.4
Visceral metastases		
Yes	121	87.1
No	18	12.9
Number of metastatic sites		
1	17	12.2
2–3	70	50.4
≥ 4	52	37.4

Abbreviations: *HR* hormone receptor positive, *HER2* human epidermal growth factor receptor 2

[a] The number of breast cancer patients

Analysis of clonal evolution

Clonal structure analysis of ctDNA samples collected at different times revealed distinct evolutionary patterns. Patient P1 exhibited a "linear" evolution pattern, where all progeny clones derived from a single progenitor clone evolved sequentially to form new clones with survival advantages (Fig. 2A). In contrast, patient P2 displayed a "branched" evolution pattern, where two distinct branching clones emerged from the progenitor clone and evolved independently into new subclones due to antitumor treatment and external selection (Fig. 2B). In other words, a linear evolution pattern is characterized by each node in the clonal evolution tree having only one child clone, indicating a straightforward and unbranched development. Conversely, a branched evolution pattern is indicated by a node with at least two child clones, signifying multiple branches in clonal development. In this study, 40 patients (31.0%) exhibited linear evolution patterns, while 89 patients (69.0%) exhibited branched evolution patterns. These findings suggest that most breast cancer patients undergo branched evolution after multiple lines of treatment for recurrence and metastasis under antitumor treatment and external selection. Specifically, branched evolution was more common in HR-positive patients and HER2-positive breast cancers, whereas linear evolution was more common in triple-negative breast cancer (Table 2).

Relationships between clonal evolution patterns and treatment efficacy and prognosis

An analysis was conducted on 86 patients for whom PFS data were available after the second ctDNA test. Of these, 22 received chemotherapy, 7 received endocrine

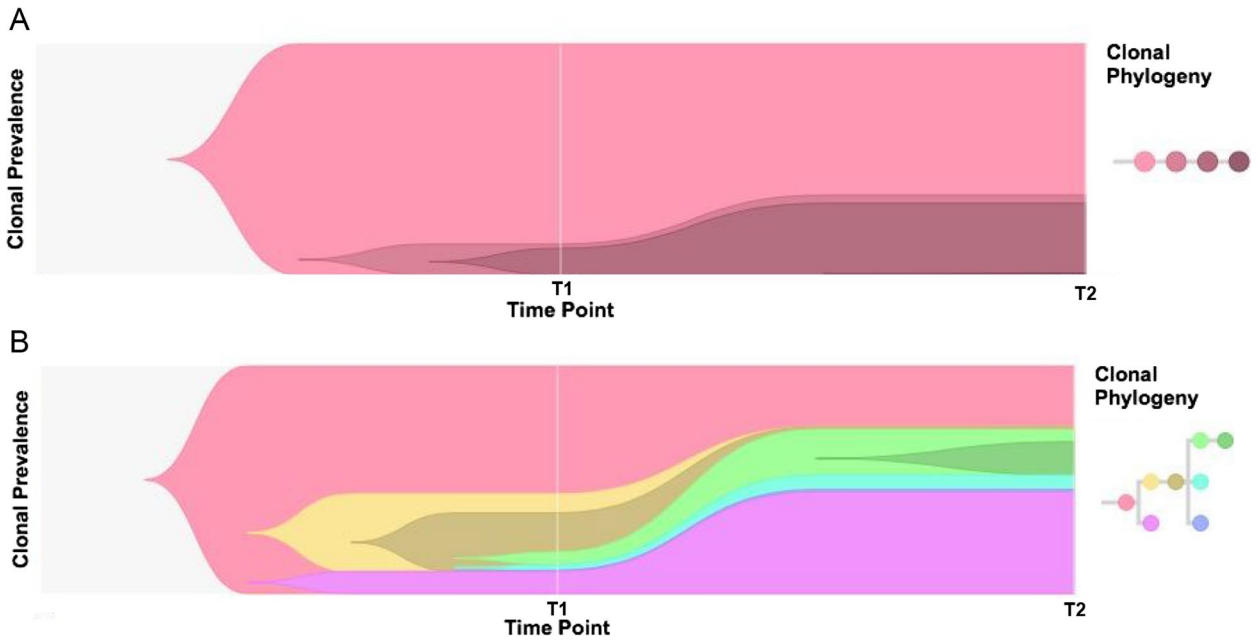

Fig. 2 Tumor clonal evolution patterns. **A** Linear evolution. **B** Branched evolution

Table 2 Evolutionary patterns in different molecular subtypes of breast cancer patients

Tumor evolutionary pattern	Molecular subtype/number of cases (%)		
	HR + /HER2 −	HER2 +	HR − /HER2 −
Linear evolution	24 (30.0)	8 (22.2)	8 (61.5)
Branched evolution	56 (70.0)	28 (77.8)	5 (38.5)

Abbreviations: *HR* hormone receptor, *HER2* human epidermal growth factor receptor 2

therapy, 34 received chemotherapy combined with targeted therapy, 18 received endocrine therapy combined with targeted therapy, and 5 received other treatments (such as targeted therapy or radiotherapy). Among these patients, 24 exhibited linear evolution, while 62 exhibited branched evolution. Patients with branched evolution had a median PFS of 4.0 months compared to 2.2 months for those with linear evolution (hazard ratio (HR), 0.52; 95% confidence interval (CI), 0.30–0.91; $P=0.004$) (Fig. 3A). Next, we conducted further analysis on patients received chemotherapy and other treatments after the second ctDNA test. Among the 56 patients treated with chemotherapy, 17 exhibited linear evolution, while 39 exhibited branched evolution. Similarly, patients with branched evolution exhibited better prognosis, with median PFS extended by 2.3 months compared to those with linear evolution (HR, 0.48; 95% CI, 0.23–0.95; $P=0.006$). Among the remaining 30 patients who received other treatments, 7 exhibited linear evolution, while 23 exhibited branched evolution. The median PFS for patients with branched and linear evolution patterns were 3.4 months and 2.3 months, respectively. However, the difference was not statistically significant (HR, 0.74; 95% CI, 0.29–1.87; $P=0.490$) (Additional file 2: Fig. S3A–B).

To adjust for clinical factors, we included molecular subtype ($P=0.010$) and clonal evolution patterns in the Cox regression model. Results indicated that patients with branched evolution had a lower likelihood of disease progression (HR, 0.53; 95% CI, 0.32–0.87; $P=0.012$). For OS in 129 patients, those with linear evolution patterns had a median OS of 56.3 months. The survival data for patients with branched evolution were not yet mature, but a trend towards better prognosis was observed in these patients compared to those with linear evolution (HR, 0.55; 95% CI, 0.28–1.09; $P=0.052$) (Fig. 3B). Cox regression analysis for OS included molecular subtype ($P=0.496$), number of metastatic sites ($P=0.001$) and clonal evolution patterns. The results showed that patients with less metastases had a better OS (HR, 0.38; 95% CI, 0.22–0.64; $P<0.001$), and the evolution pattern was not significantly related to OS.

The role of TP53 and PIK3CA in tumor heterogeneity

We analyzed the roles of the two most frequently mutated genes, TP53 and PIK3CA, in tumor heterogeneity. Trunk-resistant mutations refer to gene mutations that occur in clones at the early stage of tumor formation, while nontrunk-resistant mutations are those that appear

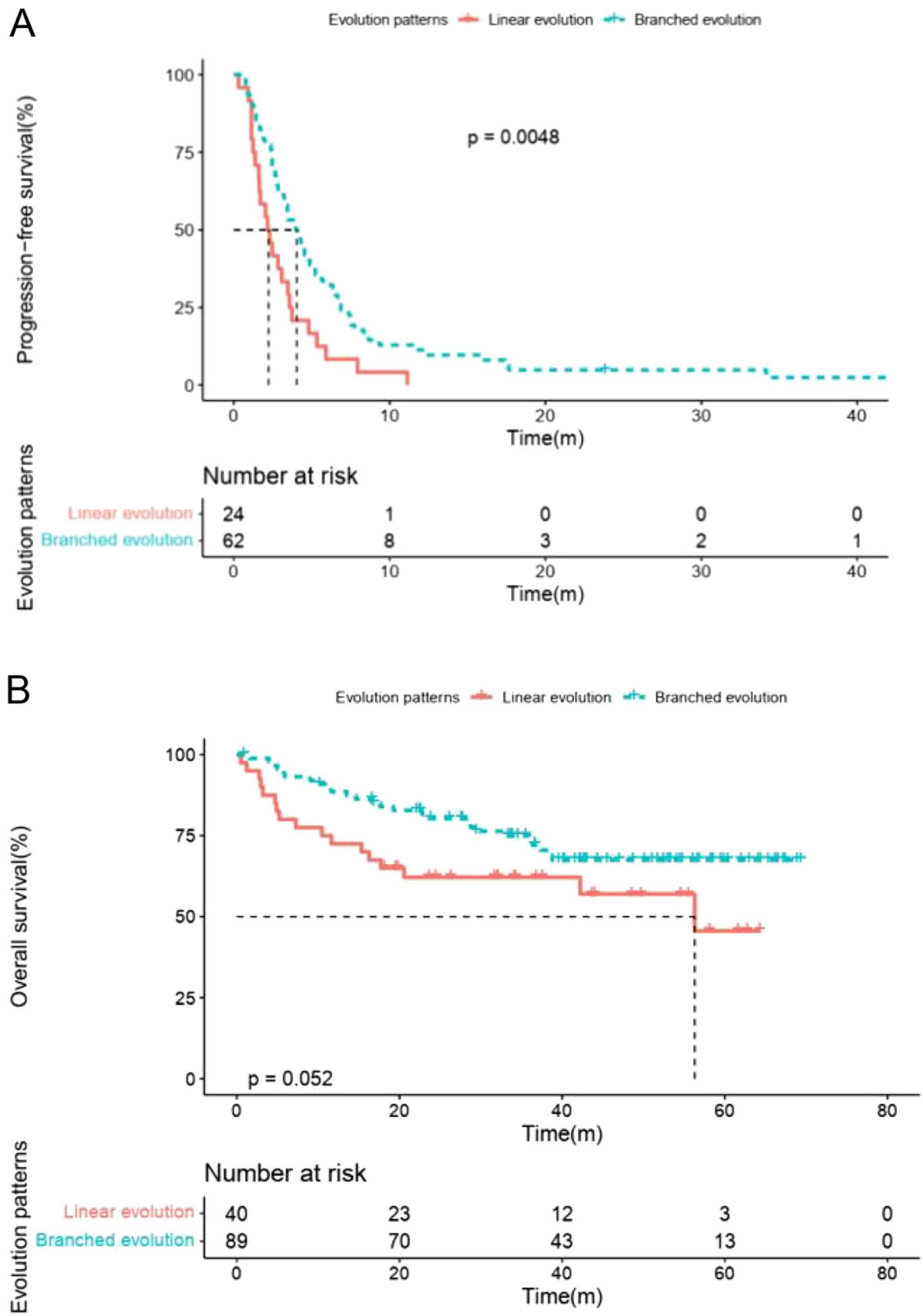

Fig. 3 Survival analysis based on the clonal evolution patterns. **A** PFS analysis between patients with linear and branched evolution. **B** OS analysis between patients with linear and branched evolution

in clones during tumor evolution. We analyzed the PFS data of 86 patients after the second ctDNA test. Twenty patients carried TP53 trunk-resistant mutations, while 31 patients carried PIK3CA trunk-resistant mutations. The median PFS for patients with TP53 trunk-resistant mutations was 3.3 months compared to 3.4 months for those without trunk-resistant mutations (HR, 1.09; 95% CI, 0.65–1.83; $P=0.722$) (Additional file 2: Fig. S4A). Patients with PIK3CA trunk-resistant mutations had a median PFS of 2.5 months, compared to 4.0 months for those without these mutations, indicating a trend towards more aggressive tumor progression, though not statistically significant (HR, 1.48; 95% CI, 0.92–2.38; $P=0.076$) (Additional file 2: Fig. S4B).

Survival analysis of 46 patients with available OS data revealed that 13 patients carried TP53 trunk-resistant mutations, and 15 patients carried PIK3CA trunk-resistant mutations. Patients with TP53 trunk-resistant mutations had a median OS of 11.6 months, compared to 16.3 months for those without these mutations (HR, 1.59; 95% CI, 0.77–3.28; $P=0.144$) (Additional file 2: Fig. S4C). For patients with PIK3CA trunk-resistant mutations, the median OS was 16.3 months, compared to 11.6 months for those without these mutations (HR, 1.08; 95% CI, 0.58–1.96; $P=0.823$) (Additional file 2: Fig. S4D).

TER as a predictor of treatment efficacy and prognosis in the primary cohort

Using the bootstrap resampling method, we identified the optimal threshold values for the TER corresponding to OS and found that the optimal threshold value for the TER was $14.15^{*}10^{-4}$ in the primary cohort. An analysis of the clinical and pathological factors influencing the TER revealed that the TER was not significantly correlated with age at diagnosis, pathological type, hormone receptor status, HER2 status, molecular subtype, the presence of visceral metastasis, or the number of metastatic organs ($P>0.05$) (Additional file 3).

Survival analysis revealed that the median PFS in the low TER (TER-L) group was longer than that in the high TER (TER-H) group for the next treatment regimen (4.3 months *vs.* 2.8 months, (HR, 0.64; 95% CI, 0.41–0.98; $P=0.031$)) (Fig. 4A). The analysis of PFS indicated that HR-positive patients and HER2-positive patients had better PFS than TNBC patients ($P=0.004$). Including TER, molecular subtype, number of metastatic organs ($P=0.604$), and next-line treatment regimen ($P=0.838$) in the Cox regression model, we found that TER was an influencing factor for PFS, with TER-L status associated with better PFS (HR, 0.62; 95% CI, 0.40–0.96; $P=0.033$) (Table 3).

Further exploration of the optimal treatment strategies for TER-L and TER-H patients is needed. In the TER-L group, the median PFS for patients receiving chemotherapy with or without targeted therapy and those receiving endocrine therapy with or without targeted therapy were 4.2 months and 5.9 months, respectively (HR, 1.05; 95% CI, 0.54–2.04; $P=0.930$) (Additional file 2: Fig. S5A). In the TER-H group, no statistically significant difference was observed in median PFS between the two treatment modalities (both 2.8 months; (HR, 0.90; 95% CI, 0.45–1.81; $P=0.760$)) (Additional file 2: Fig. S5B).

Analysis of 46 patients' mortality events showed that the median OS for the TER-L group was 21.4 months compared to 11.5 months for the TER-H group, indicating significantly longer OS in the TER-L group (HR, 0.49; 95% CI, 0.27–0.88; $P=0.009$; Fig. 4B). Univariate analysis of OS did not include visceral metastasis due to only one patient without visceral metastasis. Age ($P=0.299$), molecular subtype ($P=0.202$), and number of metastatic organs ($P=0.641$) did not significantly differ between groups. Including the number of metastatic organs and molecular subtype in the Cox regression model, we found that TER was an influencing factor for OS, with the TER-L group having a better prognosis (HR, 0.45; 95% CI, 0.24–0.85; $P=0.013$) (Table 4).

Validation of the predictive performance of TER and clonal evolution patterns

To validate the predictive value of the TER, we expanded the sample size and collected data from 267 metastatic breast cancer patients (median age at diagnosis, 40 years (range, 24–74 years)) from multiple hospitals nationwide. The validation cohort consisted of 58 patients in the TER-H subgroup and 209 patients in the TER-L subgroup. Although the median OS was not reached in either group, patients in the TER-L group had a significantly better prognosis than those in the TER-H group (HR, 0.41; 95% CI, 0.21–0.83; $P<0.001$). These results demonstrate that the TER has good prognostic efficacy (Fig. 5A). In addition, we also validated the relationship between evolution patterns and OS in the validation cohort and found that no association between clonal evolution patterns and survival ($P=0.970$; Fig. 5B).

Discussion

Recent advances in early diagnosis and targeted therapy development have significantly improved the diagnosis and treatment of breast cancer [20]. However, patients with metastatic breast cancer still face a high risk of mortality [21]. Previous research has shown that tumor heterogeneity and clonal evolution are associated with treatment resistance and disease progression in metastatic breast cancer [22]. Despite this, there is limited research on genomic dynamics and clonal structural changes in these patients. Given the high heterogeneity

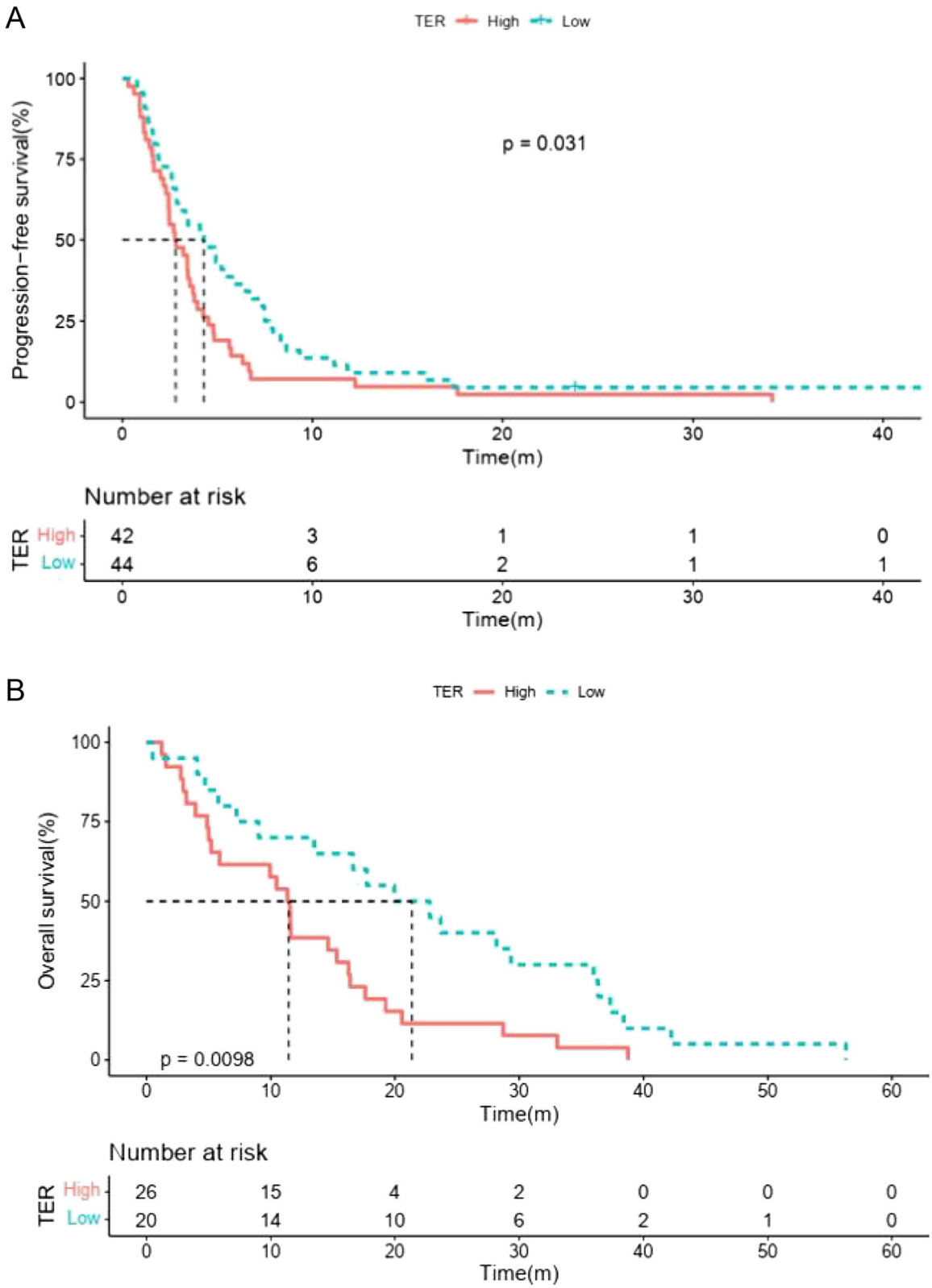

Fig. 4 Survival analysis based on the TER in the training cohort. **A** PFS analysis based on the TER in the training cohort. **B** OS analysis based on the TER in the training cohort

Table 3 Cox regression analyses between PFS and clinical characteristics ($n=86$)

Variables	Univariate analysis		Multivariable analysis	
	hazard ratio (95% CI)	*P*	hazard ratio (95% CI)	*P*
TER Low vs. high	0.64 (0.41–0.98)	0.031	0.62 (0.40–0.96)	0.033
Molecular subtype TNBC vs. HR + and HER2 +	3.82 (0.71–20.7)	0.004	-	-
Number of metastatic organs ≥ 4 vs. 2–3 and 1	1.24 (0.78–1.95)	0.604	-	-
Next-line treatment regimen Chemotherapy vs. other regimens	1.05 (0.67–1.65)	0.838	-	-

Abbreviations: *HR* hormone receptor, *HER2* human epidermal growth factor receptor 2, *TNBC* triple-negative breast cancer, *TER* tumor clonal evolution rate, *CI* confidence interval

Table 4 Cox regression analyses between OS and clinical characteristics ($n=46$)

Variables	Univariate analysis		Multivariable analysis	
	Hazard ratio (95% CI)	*P*	Hazard ratio (95% CI)	*P*
TER Low vs. high	0.49 (0.27–0.88)	0.009	0.45 (0.24–0.85)	0.013
Molecular subtype TNBC vs. HR + and HER2 +	2.41 (0.53–10.96)	0.202	1.21 (0.89–1.64)	0.232
Number of metastatic organs ≥ 4 vs. 2–3, 1	1.15 (0.63–2.06)	0.641	1.14 (0.66–1.97)	0.628

Abbreviations: *HR* hormone receptor, *HER2* human epidermal growth factor receptor 2, *TNBC* triple-negative breast cancer, *TER* tumor clonal evolution rate, *CI* confidence interval

of breast cancer, the molecular subtypes and genetic mutation characteristics in metastatic lesions may differ significantly from those in the primary tumor [23]. Therefore, repeat biopsies of metastatic lesions are clinically significant for understanding changes in molecular subtypes following recurrence and metastasis [24, 25]. However, obtaining tumor tissue from metastatic lesions can be challenging and invasive, with high risks. Moreover, a single tissue biopsy may not capture all biological characteristics due to the spatial heterogeneity of tumors. ctDNA, shed from circulating tumor cells into the peripheral blood, provides real-time, dynamic information on tumor heterogeneity and overcomes the limitations of traditional biopsies. This approach has advantages such as minimally invasive sampling and easier accessibility [5, 6]. Previous studies have shown that peripheral blood ctDNA has gene sequences that are highly similar to those of paired tumor tissues. This similarity enables the inference of tumor cell clonal structures through the analysis of mutation spectra and frequencies in ctDNA [6, 26, 27].

This study included patients with metastatic breast cancer who had undergone multiple treatments. Blood samples from various time points were subjected to ctDNA sequencing targeting 1021 genes. In the primary cohort, ctDNA mutations were detected in 84.5% of patients at baseline, which is consistent with previous reports [28]. In the second ctDNA sequencing, the detection rate of ctDNA mutations increased to 89.2%, suggesting an increase in tumor burden with disease progression. Analysis of the results from the two ctDNA tests revealed TP53, PIK3CA, ERBB2, and ESR1 as the top four genes according to mutation frequency. Mutation frequencies were greater in the second sample than in the first, which is consistent with data on commonly reported genes associated with breast cancer recurrence and metastasis [17].

Multiomic sequencing has revealed key aspects of cancer evolution in a variety of solid cancers [6, 29–31]. In our study, PyClone and CITUP software were used to analyze tumor-associated gene mutations and determine the process of tumor clonal evolution. Most patients exhibited adaptive alterations under the pressure of a series of anticancer treatments, resulting in new mutational subclones and a branched clonal evolutionary pattern. A minority of patients exhibited linear evolution. Similar features were observed in patients with postoperative glioma recurrence, where a significant increase in new mutations was observed following tumor recurrence and metastasis compared to the primary lesion, presenting a branched clone evolutionary pattern [32]. Furthermore, we found that patients with linear evolution patterns were more prone to disease progression during treatment. This result was also observed in the chemotherapy-treated subgroup, where patients with branched evolution had a median PFS 2.3 months longer than

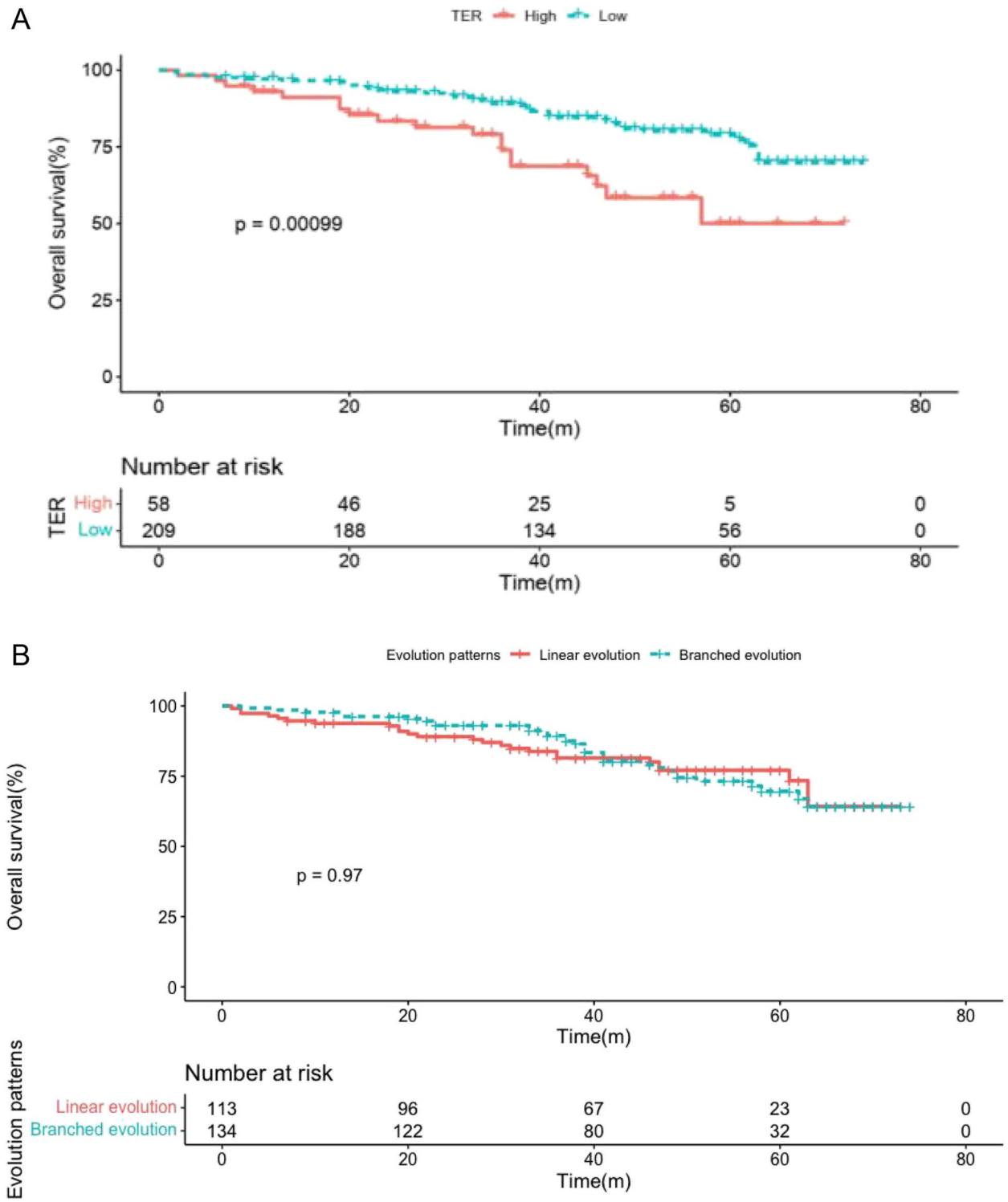

Fig. 5 OS analysis based on the TER and clonal evolution patterns in the validation cohort. **A** OS analysis between TER-H and TER-L in validation patients. **B** OS analysis between patients with linear and branched evolution in the validation cohort

those with linear evolution. However, among patients receiving other treatments, we observed that those with branched evolution had a median PFS of 1.1 months longer than those with linear evolution, although this difference was not statistically significant. We speculate that the reason for this result may be related to the smaller sample size of patients in the subgroup. Most HR-positive patients and HER2-positive breast cancer patients exhibited branched evolution, while most TNBC patients demonstrated linear evolution, which was consistent with the poor prognosis associated with TNBC. Possible reasons include that in linear evolutionary models, some primary-driven tumor-associated genes play an active role in inducing treatment resistance. But among clear cell renal cell carcinoma patients, those with branching evolution were more aggressive than those with linear evolution [30]. This difference may be related to intertumor heterogeneity.

Our study revealed that monitoring the dynamic changes in the clonal evolution of tumors in patients can reflect the effectiveness of anticancer treatments. We introduce TER as a novel concept that integrates the clonal evolution characteristics detected between two ctDNA tests, along with the impact of both the primary clone and the subclones generated under treatment pressure on therapy. Compared to clonal evolution patterns, the TER reflects the dynamic evolution speed of tumor clones. We determined the optimal threshold for dividing patients into TER-L and TER-H groups and found that patients in the TER-L group had better median PFS and OS than those in the TER-H group. This suggests that patients with slower clonal evolution experienced longer PFS and OS with subsequent lines of anticancer treatment. Validation with a larger cohort from multiple centers confirmed that patients in the TER-L group had a better prognosis, indicating that TER could serve as a biomarker for predicting treatment efficacy and prognosis.

Previous studies have frequently utilized the tumor mutational burden (TMB) to reflect the instability and mutation load of tumors. Currently, the TMB is utilized as a biomarker for predicting the response to immune checkpoint inhibitor therapy. However, the role of the TMB in the prediction of prognosis remains unclear [33]. A study revealed that a high TMB in colorectal cancer and esophageal cancer indicates a poor prognosis, while a high TMB in tumors such as urothelial carcinoma, endometrial carcinoma, and gastric adenocarcinoma is associated with a better prognosis [34]. Additionally, McGrail et al. reported that TMB in breast cancer is not linked to patient prognosis [35]. Preclinical studies have shown that tumors with high TMB and low intratumoral heterogeneity exhibit slow growth and low invasiveness in both in vivo and in vitro experiments [36]. This suggests that the use of the TMB to monitor treatment efficacy in some tumors may have limitations and inconsistencies. In this study, we considered the clonal evolution characteristics of tumor-related gene mutations and tumor heterogeneity in terms of time and space. We proposed the TER as a new indicator to reflect the rate of change in the tumor evolution process and tumor heterogeneity. Our findings suggest that TER can predict treatment efficacy and prognosis, overcoming TMB's limitations in prognosis prediction. Thus, TER may serve as a potential biomarker, providing new evidence for using ctDNA as a molecular marker in predicting the efficacy and prognosis of breast cancer treatment.

This study also has several limitations. First, it is a retrospective study, and more prospective clinical studies are needed to confirm the role of the TER in predicting tumor efficacy and prognosis. Second, all patients included in this study had breast cancer, and further exploration of the clinical application value of the TER in other types of tumors is needed.

Conclusions

This study analyzed peripheral blood ctDNA in patients with metastatic breast cancer to elucidate genetic changes and tumor clonal evolution following multiple treatments. It provides a deeper understanding of tumor evolution in response to anticancer therapies. Patients exhibiting branched evolution pattern experienced poor outcomes. The development of the TER as a novel molecular indicator for predicting treatment efficacy and prognosis supports the use of ctDNA as a valuable molecular marker in breast cancer treatment.

Abbreviations

VAF	Variant allele frequencies
ctDNA	Circulating tumor DNA
CT	Computed tomography
CI	Confidence interval
HR	Hormone receptor
HER2	Human epidermal growth factor receptor 2
MRD	Minimal residual disease
OS	Overall survival
PFS	Progression-free survival
RECIST	Response Evaluation Criteria in Solid Tumors
TER	Tumor clonal evolution rate
TNBC	Triple-negative breast cancer
TMB	Tumor mutational burden

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12916-025-03959-6.

Additional file 1: Genes targeted for capture sequencing in the 1021 gene panel.

Additional file 2: Figure S1. Mutational characteristics at baseline and after disease progression following treatment regimens. A Mutational characteristic at baseline. B Mutational characteristic after disease progression

following treatment regimens. Figure S2. Comparison of two ctDNA tests for high frequency genes. Figure S3. The relationship between different treatments and clonal evolution patterns. A Patients who received chemotherapy and chemotherapy combined with targeted therapy. B Patients who received treatment without chemotherapy. Figure S4. Survival analysis based on characterization of TP53 and PIK3CA resistance mutations. A PFS analysis between patients with and without TP53 trunk resistant mutations. B PFS analysis between patients with and without PIK3CA trunk resistant mutations. C OS analysis between patients with and without TP53 trunk resistant mutations. D OS analysis between patients with and without PIK3CA trunk resistant mutations. Figure S5. Differences in efficacy of chemotherapy ± targeted therapy compared to endocrine therapy ± targeted therapy in the TER-L and TER-H patients. A TER-L group patients. B TER-H group patients.

Additional file 3: Comparison of baseline characteristics between TER-H and TER-L patients.

Acknowledgements
Not applicable.

Authors' contributions
The concept and design of the study were contributed by Lv, Lan, Q. Guo, Chen, and Ma. All authors were involved in the acquisition, analysis, or interpretation of data, as well as in the revision of the manuscript . Lv, Lan, and Q. Guo conducted the statistical analysis. Funding was obtained by Lan and Ma. All authors provided administrative, technical, or material support, and supervision was carried out by Lv, Lan, Q. Guo, Chen, and Ma. All authors reviewed the manuscript.

Funding
This work was supported by The National Key Research and Development Program of China (2021YFF1201300) and the CAMS Innovation Fund for Medical Sciences (CIFMS) (2021-I2M-1–014 and 2023-I2M-C&T-B-081).

Data availability
The partial genome sequence data for the primary cohort in this study could be obtained from China National GeneBank DataBase (CNGBdb) under the following accession: https://db.cngb.org/search/project/CNP0001305/. Due to institutional regulations, the sequencing data for the validation cohort cannot be uploaded to a public repository, it is available upon reasonable request from the corresponding author.

Declarations

Ethics approval and consent to participate
The study was approved by the Institutional Review Board of the National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital (Approval No.16–038/1117). All patients provided written informed consent to participate in the biomarker analysis study.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Medical Oncology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100021, China. [2]Department of Computer Science and Technology & Institute of Artificial Intelligence & BNRist, Tsinghua University, Beijing 100084, China. [3]State Key Laboratory of Molecular Oncology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100021, China. [4]Geneplus-Beijing Institute, Beijing 102206, China.

Received: 11 July 2024 Accepted: 18 February 2025
Published online: 25 February 2025

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021;71(3):209–49.
2. Gupta GP, Massagué J. Cancer metastasis: building a framework. Cell. 2006;127(4):679–95.
3. Valastyan S, Weinberg RA. Tumor metastasis: molecular insights and evolving paradigms. Cell. 2011;147(2):275–92.
4. Ding L, Ellis MJ, Li S, Larson DE, Chen K, Wallis JW, et al. Genome remodelling in a basal-like breast cancer metastasis and xenograft. Nature. 2010;464(7291):999–1005.
5. Beca F, Polyak K. Intratumor heterogeneity in breast cancer. Adv Exp Med Biol. 2016;882:169–89.
6. Gerlinger M, Rowan AJ, Horswell S, Math M, Larkin J, Endesfelder D, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. N Engl J Med. 2012;366(10):883–92.
7. Lee J, Kim ST, Kim K, Lee H, Kozarewa I, Mortimer PGS, et al. Tumor genomic profiling guides patients with metastatic gastric cancer to targeted treatment: the VIKTORY umbrella trial. Cancer Discov. 2019;9(10):1388–405.
8. Dawson SJ, Tsui DW, Murtaza M, Biggs H, Rueda OM, Chin SF, et al. Analysis of circulating tumor DNA to monitor metastatic breast cancer. N Engl J Med. 2013;368(13):1199–209.
9. Yi Z, Liu B, Sun X, Rong G, Wang W, Li H, et al. Safety and efficacy of sirolimus combined with endocrine therapy in patients with advanced hormone receptor-positive breast cancer and the exploration of biomarkers. Breast. 2020;52:17–22.
10. Hu ZY, Xie N, Tian C, Yang X, Liu L, Li J, et al. Identifying circulating tumor DNA mutation profiles in metastatic breast cancer patients with multiline resistance. EBioMedicine. 2018;32:111–8.
11. Liu B, Hu Z, Ran J, Xie N, Tian C, Tang Y, et al. The circulating tumor DNA (ctDNA) alteration level predicts therapeutic response in metastatic breast cancer: novel prognostic indexes based on ctDNA. Breast. 2022;65:116–23.
12. Pellini B, Chaudhuri AA. Circulating tumor DNA minimal residual disease detection of non-small-cell lung cancer treated with curative intent. J Clin Oncol. 2022;40(6):567–75.
13. Moding EJ, Nabet BY, Alizadeh AA, Diehn M. Detecting liquid remnants of solid tumors: circulating tumor DNA minimal residual disease. Cancer Discov. 2021;11(12):2968–86.
14. Garcia-Murillas I, Schiavon G, Weigelt B, Ng C, Hrebien S, Cutts RJ, et al. Mutation tracking in circulating tumor DNA predicts relapse in early breast cancer. Sci Transl Med. 2015;7(302):302ra133.
15. Olsson E, Winter C, George A, Chen Y, Howlin J, Tang MH, et al. Serial monitoring of circulating tumor DNA in patients with primary breast cancer for detection of occult metastatic disease. EMBO Mol Med. 2015;7(8):1034–47.
16. Ma F, Guan Y, Yi Z, Chang L, Li Q, Chen S, et al. Assessing tumor heterogeneity using ctDNA to predict and monitor therapeutic response in metastatic breast cancer. Int J Cancer. 2020;146(5):1359–68.
17. Yi Z, Ma F, Rong G, Liu B, Guan Y, Li J, et al. The molecular tumor burden index as a response evaluation criterion in breast cancer. Signal Transduct Target Ther. 2021;6(1):251.
18. Roth A, Khattra J, Yap D, Wan A, Laks E, Biele J, et al. PyClone: statistical inference of clonal population structure in cancer. Nat Methods. 2014;11(4):396–8.
19. Malikic S, McPherson AW, Donmez N, Sahinalp CS. Clonality inference in multiple tumor samples using phylogeny. Bioinformatics. 2015;31(9):1349–56.
20. Harbeck N, Penault-Llorca F, Cortes J, Gnant M, Houssami N, Poortmans P, et al. Breast cancer. Nat Rev Dis Primers. 2019;5(1):66.
21. Tellez-Gabriel M, Knutsen E, Perander M. Current status of circulating tumor cells, circulating tumor DNA, and exosomes in breast cancer liquid biopsies. Int J Mol Sci. 2020;21(24):9457.

22. Cresswell GD, Nichol D, Spiteri I, Tari H, Zapata L, Heide T, et al. Mapping the breast cancer metastatic cascade onto ctDNA using genetic and epigenetic clonal tracking. Nat Commun. 2020;11(1):1446.
23. Yi ZB, Yu P, Zhang S, Wang WN, Han YQ, Ouyang QC, et al. Profile and outcome of receptor conversion in breast cancer metastases: a nation-wide multicenter epidemiological study. Int J Cancer. 2021;148(3):692–701.
24. Carlson RW, Allred DC, Anderson BO, Burstein HJ, Edge SB, Farrar WB, et al. Metastatic breast cancer, version 1.2012: featured updates to the NCCN guidelines. J Natl Compr Canc Netw. 2012;10(7):821–9.
25. Van Poznak C, Somerfield MR, Bast RC, Cristofanilli M, Goetz MP, Gonzalez-Angulo AM, et al. Use of biomarkers to guide decisions on systemic therapy for women with metastatic breast cancer: American Society of Clinical Oncology Clinical Practice Guideline. J Clin Oncol. 2015;33(24):2695–704.
26. Ai X, Cui J, Zhang J, Chen R, Lin W, Xie C, et al. Clonal architecture of EGFR mutation predicts the efficacy of EGFR-Tyrosine kinase inhibitors in advanced NSCLC: a prospective multicenter study (NCT03059641). Clin Cancer Res. 2021;27(3):704–12.
27. Nong J, Gong Y, Guan Y, Yi X, Yi Y, Chang L, et al. Circulating tumor DNA analysis depicts subclonal architecture and genomic evolution of small cell lung cancer. Nat Commun. 2018;9(1):3114.
28. Rong G, Yi Z, Ma F, Guan Y, Xu Y, Li L, et al. DNA damage response as a prognostic indicator in metastatic breast cancer via mutational analysis. Ann Transl Med. 2021;9(3):220.
29. Zhai W, Lim TK, Zhang T, Phang ST, Tiang Z, Guan P, et al. The spatial organization of intra-tumour heterogeneity and evolutionary trajectories of metastases in hepatocellular carcinoma. Nat Commun. 2017;8:4565.
30. Laruelle A, Manini C, López JI, Rocha A. Early evolution in cancer: a mathematical support for pathological and genomic evidence in clear cell renal cell carcinoma. Cancers. 2023;15(24):5897.
31. Kang M, Na HY, Ahn S, Kim JW, Lee S, Ahn S, et al. Gallbladder adenocarcinomas undergo subclonal diversification and selection from precancerous lesions to metastatic tumors. ELife. 2022;11:11.
32. Johnson BE, Mazor T, Hong C, Barnes M, Aihara K, McLean CY, et al. Mutational analysis reveals the origin and therapy-driven evolution of recurrent glioma. Science. 2014;343(6167):189–93.
33. Addeo A, Friedlaender A, Banna GL, Weiss GJ. TMB or not TMB as a biomarker: that is the question. Crit Rev Oncol Hematol. 2021;163:103374.
34. Wu HX, Wang ZX, Zhao Q, Chen DL, He MM, Yang LP, et al. Tumor mutational and indel burden: a systematic pan-cancer evaluation as prognostic biomarkers. Ann Transl Med. 2019;7(22):640.
35. McGrail DJ, Pilié PG, Rashid NU, Voorwerk L, Slagter M, Kok M, et al. High tumor mutation burden fails to predict immune checkpoint blockade response across all cancer types. Ann Oncol. 2021;32(5):661–72.
36. Wolf Y, Bartok O, Patkar S, Eli GB, Cohen S, Litchfield K, et al. UVB-induced tumor heterogeneity diminishes immune response in melanoma. Cell. 2019;179(1):219-35.e21.

Publisher's Note