Breast Cancer Research and Treatment (2019) 178:63–73
https://doi.org/10.1007/s10549-019-05375-w

PRECLINICAL STUDY

Plasma cell-free DNA chromosomal instability analysis by low-pass whole-genome sequencing to monitor breast cancer relapse

Huanhuan Zhou[1] · Xiao-Jia Wang[1] · Xiyi Jiang[2] · Ziliang Qian[3] · Tianhui Chen[2] · Yue Hu[4] · Zhan-Hong Chen[1] · Yun Gao[1] · Rong Wang[1] · Wei-Wu Ye[1] · Wen-Ming Cao[1]

Received: 31 January 2019 / Accepted: 23 July 2019 / Published online: 30 July 2019

Abstract

Background Chromosomal instabilities (CIN) of plasma cell-free DNA (cfDNA) are common in breast cancer. We aimed to investigate the value of cfDNA CIN in monitoring the breast cancer relapse and additionally to compare it with the traditional biomarkers (CA15-3 and CEA).

Methods Overall 62 recurrent breast cancer patients and 20 healthy controls were recruited. Low-pass whole-genome sequencing (LPWGS) was performed to detect cfDNA CIN. A CIN score was calculated. The performance of CA15-3, CEA, and CIN score in monitoring the recurrence was investigated with receiver operating characteristic (ROC) curve and the area under curve (AUC). Multivariable Cox proportional hazard model was established to analyze the correlations between copy number gain/loss and disease-free survival (DFS).

Results cfDNA CIN achieved the positive rate of 77.6% [(95% confidence interval (CI) 73.4–95.3%)] among recurrent breast cancer patients, with an AUC value of 0.933, superior to CA15-3 (positive rate: 38.7%; AUC: 0.864) and CEA (positive rate: 41.93%; AUC: 0.878) ($P < 0.01$). The combination of cfDNA CIN with two biomarkers further increased the positive rate to 88.7% (95% confidence interval 77.5–95.0%). cfDNA CIN achieved better performance in patients with shorter DFS (≤ 41 months), with an AUC value of 0.975.

Conclusions cfDNA CIN yields a higher accuracy in monitoring breast cancer recurrence compared to traditional biomarkers (CA15-3 and CEA), especially for biomarker-negative patients. The combination of cfDNA CIN to traditional biomarkers further improved the detection rate of recurrence, which may provide a new method for monitoring the early relapse of breast cancer, though further investigations are warranted.

Keywords Cell-free DNA · CIN · CA153 · CEA · Recurrent breast cancer

Huanhuan Zhou and Xiao-Jia Wang have contributed equally to this work.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10549-019-05375-w) contains supplementary material, which is available to authorized users.

✉ Wen-Ming Cao
caowm@zjcc.org.cn

[1] Department of Breast Medical Oncology, Zhejiang Cancer Hospital, Hangzhou 310022, China

[2] Group of Molecular Epidemiology & Cancer Precision Prevention, Zhejiang Academy of Medical Sciences, Hangzhou 310013, China

[3] Prophet Genomics Inc, 1229 Briarcreek ct, San Jose, CA, USA

[4] Department of Breast Surgery, The Second Affiliated Hospital, Zhejiang University, Hangzhou 310008, China

Introduction

Breast cancer is one of the most common malignancies among women in western countries and China [1, 2]. It is estimated that approximately 2.1 million breast cancer patients would be newly diagnosed worldwide in 2018, accounting for nearly 25% of the female cancer patients [1]. In China, almost 70.7 thousands breast cancer-associated deaths occurred in 2015 [2], which was predominantly attributed to the recurrence or metastasis. However, monitoring the relapse of breast cancer is an arduous task in clinical practice due to the limited sensitivity and specificity of traditional serum biomarkers.

CA15-3 is a carbohydrate-containing protein antigen called mucin 1 (MUC1), which is found to be overexpressed by MUC1 gene in patients with breast cancer [3].

It is commonly used as a blood-based tumor biomarker in monitoring the recurrences of breast cancer patients, with the sensitivity ranging from 40 to 60% [4]. Carcinoembryonic antigen (CEA) is another tumor biomarker while also shows unsatisfactory performance, with the sensitivity of 38.0%. Furthermore, the combination of CEA and CA15-3 only leads to a 2–6% improvement in sensitivity [5].

Since overexpressions of CA15-3 and CEA can also be detected in benign conditions, the false positive results may confound benign conditions with occult relapses of breast cancer and thus lead to unnecessary costs for patients [6]. For instance, a previous study reported that 5.4% (44/813) of the post-surgery patients were identified with continuous biomarker-positive (detecting CA15-3, CEA and CA125 every 6 weeks) whereas 20.5% (9/44) of them were not detected with any symptoms of relapses during the follow-up care [7].

Tumor cells release the shedding DNA fragments into bloodstream. Since the tumor DNA is a small proportion of total cell-free DNA (cfDNA), a panel of genes are captured and sequenced to extreme high coverage (such as 10000×), to detect the mutated tumor alleles. This technology has revealed promising results in the differential diagnoses of lung and breast cancer [8]. However, panel-based methods are limited by the sensitivity in cancer detection and monitoring due to the high tumor heterogeneity among patients. A panel of genes may thus be missed in some patients with gene mutations, which has not been covered by the panel as designed.

Breast cancer is a type of carcinoma with low tumor mutational burden [9]; however, chromosomal instability (CIN) is commonly to be found [10, 11]. For instance, the cancer genome atlas (TCGA) data showed that more than 93% of the breast cancer patients were detected with genome region copy gains or losses (over 5%), in which the most frequent changes were found on chromosome 1, 8, and 17 [10]. It has also been reported that copy gain on chromosome 1q was identified in about 70% of the recurrent breast cancer patients [12]. Moreover, in another study, significant cfDNA CIN were found in 96.3% of the relapsed triple-negative breast cancer patients without prior knowledge of tumor mutations [13]. The above results both indicated that cfDNA CIN might serve as a potential biomarker for monitoring the breast cancer recurrence. Currently, the technique development of low-pass whole-genome sequencing (LPWGS) makes it possible to track the chromosomal changes in circulating bloodstream through a non-invasive way, which has been applied to detect cfDNA CIN in triple-negative breast cancer [12, 13].

In this study, we aimed to investigate the value of cfDNA CIN in monitoring the recurrence of breast cancer patients. Additionally, we compared the performance of cfDNA CIN to two traditional blood-based tumor biomarkers (CA15-3 and CEA).

Methods and materials

Study population and samples

Overall 62 recurrent breast cancer patients and 20 healthy controls were admitted to Zhejiang Cancer Hospital between May 1, 2017 and May 1, 2019. All of the patients were diagnosed with stage I to IIIA breast cancer and received modified radical mastectomy or breast conserving surgery at initial treatment, and enrolled in this study when they were diagnosed as recurrent breast cancer by both imaging and pathology. None of the participant was suffered from other cancers beyond breast cancer. Blood samples were collected for cfDNA extraction and CA15-3/CEA detection from the patients when they were diagnosed as recurrent breast cancer before first-line treatment.

This study was approved by the Clinical Research Ethics Committee of Zhejiang Cancer Hospital. Written informed consents were obtained from all recruited subjects.

Data collection

The clinicopathological information of recurrent breast cancer patients including age at diagnosis, histologic type, histologic grade, molecular type, and status of estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2) were collected. Patients with ER + /PR + were considered as hormone receptor positive. CerbB-2 scores of 0/1 + and 3 + were defined as HER2-negative and HER2-positive while score of 2 + corresponding to HER2 status was further confirmed through Fluorescence in situ hybridization. Most patients were followed regularly after surgery at Zhejiang Cancer Hospital until May 1, 2019. Follow-up via telephone call was also carried out if clinical records from follow-up examinations were not available. The follow-up information including the time of follow-up, post-surgery treatment, time of recurrence and metastasis, metastatic status, and date and cause of death was collected. Disease-free survival (DFS) was defined as the time interval between the diagnosis date and the last date when vital status of the patients was available. Alive patients were censored at the date of May 1, 2019.

CA15-3 detection and CEA detection

The serum CEA and CA15-3 of breast cancer patients and 20 healthy controls were measured by CA15-3 Kits and CEA kits (SIEMENS, Inc., USA), using the cutoff value of 5 ng/ml and 30KU/L.

Next generation sequencing

Total genomic DNA and cfDNA of breast cancer patients and 20 healthy controls were isolated from plasma using the Amp Genomic DNA Kit (TIANGEN) and QIAseq cfDNA Extraction kit (Qiagen), respectively. Next generation sequencing was performed as previously described [14, 15]. DNA was fragmented into an average size of 300 bp (cfDNA without fragmentation), and then 100 ng of fragmented genomic DNA (cfDNA 10 ng) was used for the construction of sequencing library (NEBnext Ultra II). Then, 8 bp barcoded sequencing adaptors were ligated with DNA fragments and amplified by PCR. Purified sequencing libraries were massively parallel sequenced by Illumina HiSeq Xten platform. Ten to twenty million reads obtained for each sample (mean coverage of 1.5X). Raw sequencing data (4G per sample) was filtered and mapped to the human reference genome version hg19. Non-uniquely mapped reads were excluded from further analysis. The human genome was then divided into 200 Kbp bins, of which the sequencing coverage for each bin was normalized (GC normalization) by a bioinformatics pipeline ultrasensitive chromosomal aneuploidy detector (UCAD). Circular binary segmentation algorithms were used to identify significant genomic breakpoints. Copy gain region was defined as a segment with ratio > 0 and size ≥ 10 M while copy loss region was defined as a segment with ratio < 0 and size ≥ 10 M. Focal amplification was defined as a statistically significant segment with normalized coverage ≥ 1.

Statistical analysis

The positive detection rates of CA15-3, CEA, and cfDNA CIN were calculated, respectively, and differences in rates of the aforementioned methods as well as the combinations were analyzed using Chi-square test or Fisher's exact test. The correlations between CA15-3, CEA, and cfDNA CIN were analyzed using Pearson correlation test, respectively.

A CIN score was further calculated according to the following formula:

$$CIN_{score} = \sum_{segment\,i} abs\left(1 - V_i\right) \times L_i$$

where V_i is the segment value and L_i is the segment length in unit of 200 Kbp.

CIN scores of ≥ 1200 and < 1200 were defined as high and low CIN, respectively. The threshold of CIN score '≥1200' was determined by using Youden Index (by using R package, 'pROC'). We used this cutoff for the best sensitivity/specificity to predict malignancies from health controls [16]. All of the healthy controls were considered as low CIN corresponding to this cutoff value.

The performance of CA15-3, CEA, and CIN score in monitoring the breast cancer recurrence was investigated with receiver operating characteristic (ROC) curve and the area under curve (AUC). DFS were estimated using the Kaplan–Meier method. Multivariable Cox proportional hazard model was established to estimate the hazard ratio (HR) analyzing the potential correlations between copy number gain/copy loss and DFS [17].

All statistical analyses were performed using SPSS version 17.0 (IBM, USA) and R version 3.4.4 (R Foundation for Statistical Computing, Vienna, Austria). A P value of < 0.05 was considered statistically significant.

Results

Clinicopathological characteristics of the recurrent breast cancer patients

The clinicopathological characteristics of recurrent breast cancer patients are listed in Table 1. Age at the time of diagnosis for patients detected with breast cancer recurrence ranged from 30–86 years. Most of the patients (79.0%) were diagnosed as invasive ductal carcinoma. Overall 87.1% (54/62) of the patients were diagnosed with luminal subtype of breast cancer, and 15 of them received hormone therapy as adjuvant treatment. Patients with viscera metastasis accounted for 72.5% (45/62), in which the most common metastasis sites were lung and liver. Moreover, positive ER, PR, and HER2 were found in 53, 48, and 6 patients, respectively.

cfDNA CIN in the recurrent breast cancer patients

Overall 16 copy gain and 18 copy loss regions were identified in the recurrent breast cancer patients (Table 2). The most frequent copy number variations included the copy gain on chromosome 1q, 7p, 8q, 12, 17q, 20 and the copy loss on chromosome 8p and 16q. Aggregated normalized coverage of each 200 k-bin is shown in Fig. 1a, which demonstrated that most of the chromosomes were detected with arm-level imbalances. For instance, copy loss and gain were found on chr8 short and long arm, respectively (Fig. 1b). Particularly, some of the chromosomes were found with copy gain or loss in whole regions such as chr20 (Fig. 1e). Copy number variations of other chromosomes are shown in supplementary material.

The most frequent copy gain genomic regions included 1q, 18, 6q12-30, 8q 7p and 20q, where potential oncogenes, IRAK1 (1q23.1) [18], MYC (8q24.21) [19], EGFR (7p11.2) [20], and BCAS1 [21] (20q13.2) were located. Copy number gains of these genomic regions were found in 36(58.1%), 34(54.8%), 34(54.8%), 29(46.8%), 19(30.6%),

Table 1 Clinicopathological characteristics of the recurrent breast cancer patients ($n=62$)

Clinicopathological characteristics	Case number (%)
Histologic type	
Invasive ductal carcinoma	49 (79.03%)
Invasive lobular carcinoma	1 (1.61%)
Other	11 (17.74%)
Unknown	1 (1.61%)
Histologic grade	
Grade I	1 (1.61%)
Grade II	28 (45.16%)
Grade III	11 (17.74%)
Unknown	22 (35.48%)
(Neo) adjuvant treatment	
Neoadjuvant and adjuvant treatment	8 (12.90%)
Adjuvant treatment alone	50 (80.65%)
None	2 (3.22%)
Unknown	2 (3.22%)
Adjuvant treatment	
Endocrine therapy alone	3 (4.84%)
Chemotherapy alone	10 (16.13%)
Chemotherapy and radiotherapy	1 (1.61%)
Endocrine therapy and radiotherapy	1 (1.61%)
Chemotherapy, endocrine therapy and radiotherapy	18 (29.03%)
Chemotherapy and endocrine therapy	17 (27.42%)
Operative type	
Modified radical mastectomy	60 (96.77%)
Breast conserving surgery	2 (3.22%)
Metastasis	
Viscera	45 (72.58%)
Non-visceral	16 (25.81%)
Unknown	1 (1.61%)
CA15-3 (KU/L)	
≥ 35	24 (38.7%)
< 35	35 (56.45%)
Unknown	3 (4.84%)
CEA (ng/ml)	
> 5	26 (41.93%)
≤ 5	33 (53.22%)
Unknown	3 (4.84%)
ER	
Positive	53 (85.48%)
Negative	7 (11.29%)
Unknown	2 (3.22%)
PR	
Positive	48 (77.42%)
Negative	11 (17.74%)
Unknown	3 (4.84%)
HER2	
Positive	6 (9.68%)
Negative	48 (77.42%)
Unknown	8 (12.90%)
Molecular type	
Luminal A	10 (16.13%)
Luminal B	44 (70.97%)
HER2 overexpression	3 (4.84%)
Triple-negative	3 (4.84%)
Unknown	2 (3.22%)

CEA carcinoembryonic antigen, *ER* estrogen receptor, *PR* progesterone receptor, *HER-2* human epidermal growth factor receptor 2

and 19(30.6%) patients, respectively. The most frequent copy loss genomic regions included 11q13.6-q25, 3p12.2-14.3, 1p, 8p, 10, 4p, 13q21-34, 9p, 4q, 6q21-23, 12p13.33-12.2, 22q11.23-q13.2, 5q31.3-35.2, 12q21.2-24.33, 22q11.23-q13.2, 6q25-27, 16q23.1-24.1, 14q24.1-24.3 and 17p13.2-11.2,where tumor suppressor genes PTEN (10q23.31) [22], CDKN2A/B (9p21.3) [23], RB1 (13q14.2) [24], and TP53 (17p13.1) [25] were located. Copy number losses of these regions were found in 52(83.9%), 52(83.9%), 48(77.4%), 48(77.4%), 44(71%), 43(69.4%), 39(62.9%), 35(56.5%), 29(46.8%), 29(46.8%), 14(22.6%), 36(58.1%), 40(64.5%), 31(50.0%), 36(58.1%), 22(35.5%), 31(50%), 38(61.3%), and 35(56.5%) patients, respectively.

Frequent focal amplifications were found on chromosome 5q31.2, 12q15, 8p11.21, 8q23.1-24.3, 11q13.3. Oncogenes including PCDHA(5q31), MDM2(12q15), IKBKB (8p11.21), OXR1 (8q23.1), and CCND1 (11q33) were found to be involved in the focal amplifications, with frequencies of 29%, 25.8%, 25.8%, 24.2% and 24.2%, respectively.

Detection results of CA15-3, CEA, and cfDNA CIN in monitoring the breast cancer recurrence

Positive CA15-3 and CEA were only found in 38.7% (24/62) and 41.9% (26/62) of the recurrent breast cancer patients, respectively; however, the positive detection rate of high cfDNA CIN (77.4%) was significantly higher than that of the above biomarkers ($P<0.01$). Moreover, the combination of CA15-3 and CEA increased the positive detection rate to 62.5%. Notably, the further combination of cfDNA CIN led to an extremely high positive detection rate of 88.7% (95% CI 77.5–94.9%) though no statistically significance was observed (Fig. 2a).

Correlations between CA15-3, CEA, and cfDNA CIN score

The correlations between CA15-3 level, CEA level, and cfDNA CIN score are shown in Fig. 2b–d. Weak linear

Table 2 Significant chromosome arm-level changes and focal amplifications identified in the recurrent breast cancer patients

Loci	Gain/loss	logP	Number (freq %)	Key genes
Arm-level changes				
1q	Gain	130.42	36 (58.1)	IRAK1
1p	Loss	130.88	48 (77.4)	
8q	Gain	212.57	29 (46.8)	MYC
10	Loss	57.63	44 (71)	PTEN
18	Gain	61.49	34 (54.8)	
7p	Gain	49.02	19 (30.6)	EGFR
7q	Gain	10.82	8 (12.9)	
13q21-34	Loss	51.50	39 (62.9)	RB1
9p	Loss	36.26	35 (56.5)	CDKN2A/B
4q	Loss	26.38	29 (46.8)	
3p12.2-14.3	Loss	37.57	52 (83.9)	
8p	Loss	39.36	48 (77.4)	
4p	Loss	27.56	43 (69.4)	
6q13-20	Gain	47.58	34 (54.8)	
20q	Gain	35.74	19 (30.6)	BCAS1
6q21-23	Loss	8.93	29 (46.8)	
11q13.6-q25	Loss	45.07	52 (83.9)	
12p13.33-12.2	Loss	12.48	14 (22.6)	
22q11.23-q13.2	Loss	22.36	36 (58.1)	
5q31.3-35.2	Loss	10.94	40 (64.5)	
12q21.2-24.33	Loss	14.59	31 (50)	
6q25-27	Loss	12.65	22 (35.5)	ESR1
17q21.31-25.3	Gain	29.91	16 (25.8)	
16q23.1-24.1	Loss	12.01	31 (50)	
14q24.1-24.3	Loss	12.55	38 (61.3)	
17p13.2-11.2	Loss	24.31	35 (56.5)	TP53
Focal amplification				
8q23.1-24.3	Gain	51.09	15 (24.2)	OXR1
8p11.21	Gain	11.09	16 (25.8)	IKBKB
8p11.23-11.22	Gain	11.31	14 (22.6)	FGFR1
12q15	Gain	18.28	16 (25.8)	MDM2
11q13.3	Gain	4.50	15 (24.2)	CCND1
11q13.5	Gain	7.38	11 (17.7)	EMSY
17q12	Gain	3.01	6 (9.7)	ERBB2
5q31.2	Gain	2.38	18 (29)	PCDHA@

Chrom chromosome, *Seg* segment, *Freq* frequency

correlations were found between CA15-3 level and cfDNA CIN score ($r=0.221$, $P=0.002$). Similarly, weak correlation was found between CEA level and cfDNA CIN score ($r=0.198$, $P=0.057$), respectively. However, CA15-3 level was found to be significantly correlated with CEA level ($r=0.586$, $P<0.001$).

The detection rate of cfDNA CIN (high/low) in patients with different CA15-3 and CEA levels was further examined. As shown in Table 3, high cfDNA CIN was found in 12 (80.0%), 7 (87.5%), 6 (54.5%), and 17 (81.0%) patients with both positive of CA15-3 and CEA, CA15-3-positive and CEA-negative, CA15-3-negative and CEA-positive, and both negative CA15-3 and CEA, respectively.

Detection results of CA15-3, CEA, and cfDNA CIN in patients with visceral and non-visceral metastasis

In previous reports [26], elevated CA15-3 was associated with visceral metastasis. As shown in Table 4, the positive rate of CA15-3 tended to be higher in patients with visceral metastasis than those with non-visceral metastasis (44.4% vs.18.8%, $P=0.05$). Similarly, no statistical significant differences were found in patients with visceral metastasis compared to those without metastasis (75.6% vs. 81.3%, $P>0.05$). However, CEA was detected to be with the opposite results, with the positive rate of 42.2% and 50.0% in patients with and without visceral metastasis, respectively ($P>0.05$). Notably, high cfDNA CIN showed the highest positive rate in patients with visceral metastasis compared to CA15-3 and CEA.

Performance of CA15-3, CEA, and CIN score to predict breast cancer relapse

The median DFS was 41 months. Patients were divided into two groups: group 1 (DFS ≤ 41 months) and group 2 (DFS > 41 months). As shown in Fig. 3a, cfDNA CIN achieved the highest AUC value of 0.933 compared to CA15-3 (0.864, $P=0.20$) and CEA (0.878, $P=0.26$) for all patients. In group 1 for patient experienced disease relapsed within 41 months, the performance of cfDNA CIN achieve AUC value of 0.975, which is higher than that of CEA (AUC = 0.866, $P=0.042$) and CA15-3(AUC = 0.895, $P=0.098$). In group 2 for patient experienced disease relapsed after 41 months, the performance of cfDNA CIN achieve AUC value of 0.897, which is similar to that of CEA (AUC = 0.894, $P=0.96$) and CA15-3(AUC = 0.827, $P=0.365$). No statistical differences were found between patients with non-visceral metastasis compared to those with metastasis, with AUC values of 0.931 and 0.932 ($P=0.91$), respectively (Fig. 3d). The predicting function of cfDNA CIN was not shown when DFS cutoff was 24, 36, and 60 months (Supplement materials Figure S1, S2, and S3).

cfDNA CIN correlated with prognosis of recurrent breast cancer patients

Positive CA15-3 at the time of breast cancer recurrence was significantly associated with shorter DFS (HR = 2.51, $P=0.002$, Fig. 4a). Similarly, high cfDNA CIN was also found to be correlated with worse DFS (HR = 1.81, $P=0.065$, Fig. 4b) though no statistically significance was

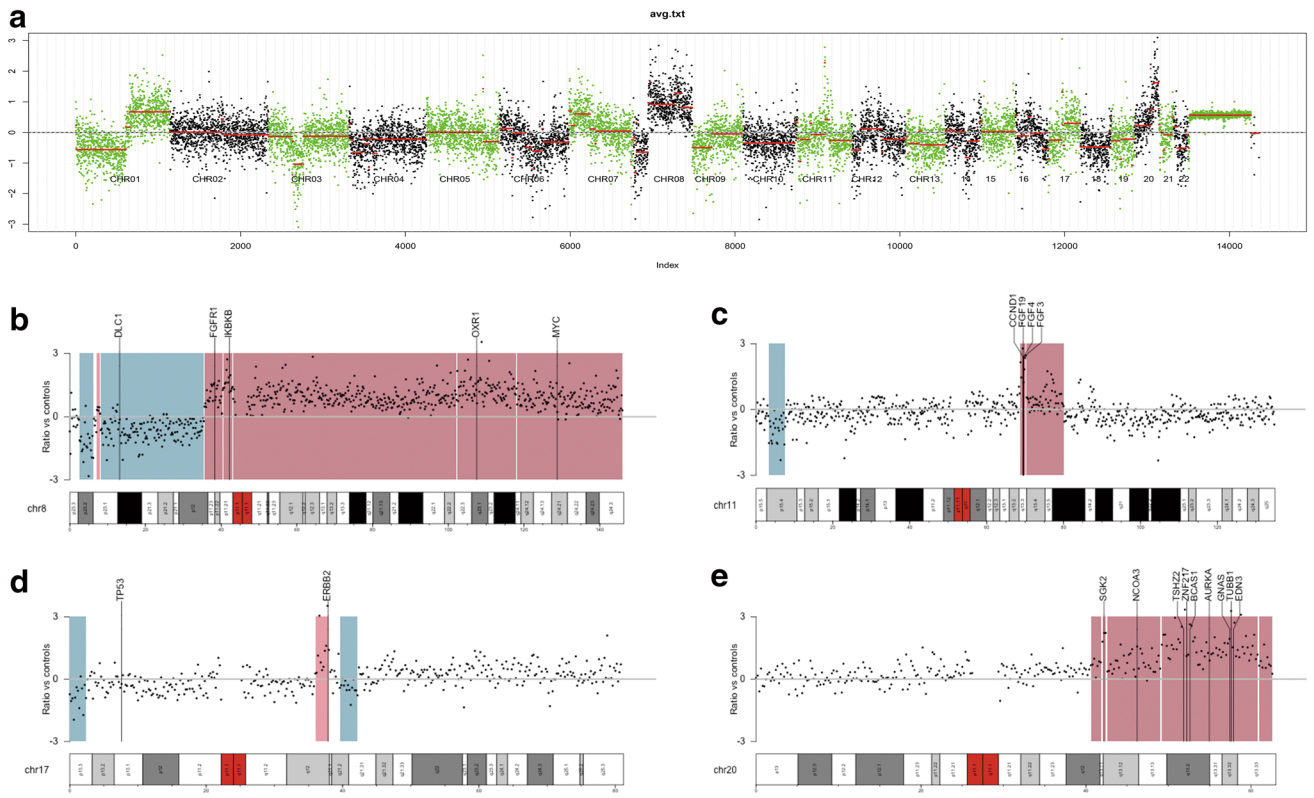

Fig. 1 Plasma cfDNA chromosomal instability (CIN) profilings of the recurrent breast cancer patients. **a** Overview of cfDNA CIN in the recurrent breast cancer patients. **b–e** Chromosomal arm-level imbalances and focal amplifications identified on chromosome 8, 11, 17, and 20, respectively

observed. Furthermore, potential correlations between the copy number variations as listed in Table 2 and DFS were also analyzed. The results showed that copy losses on Chr3p12.2-14.2 (Fig. 4c, HR = 2.49, P = 0.002) and Chr8p were significantly associated with poorer DFS (Fig. 4d, HR = 2.06, P = 0.014), with which more than half of the patients were first-time detected with breast cancer recurrence by 24-month and 36-month post-surgery, respectively.

Discussions

In this study, LPWGS assay was performed to detect cfDNA CIN. We found cfDNA CIN achieved the positive rate of 77.4% among recurrent breast cancer patients, with an AUC value of 0.933, superior to CA15-3 (positive rate: 38.7%; AUC: 0.864) and CEA (positive rate: 41.9%; AUC: 0.878). The combination of cfDNA CIN with two biomarkers further increased the positive rate to 88.7%. cfDNA CIN achieved better performance in patients with DFS ≤ 41 months compared to those with DFS > 41 months (AUC: 0.975 vs. 0.897). High cfDNA CIN was associated with shorter DFS (HR = 1.81, P = 0.065).

To our best knowledge, we reported the first study of the monitoring function of cfDNA CIN in luminal subtype breast cancer. Serum CA15-3 and CEA are two important biomarkers in monitoring the post-operative follow-up and treatment among breast cancer patients [7, 27] while was limited due to the relatively low sensitivity and specificity. In this study, only 38.7% and 41.93% of the recurrent breast cancer patients were detected with CA15-3- and CEA-positive, respectively, though the combination resulted in an improved positive rate of 62.5%. Fiorella et al. reported that CEA was elevated in 38.0% and CA15-3 in 70.2% of patients with recurrence. The combination of CEA and CA 15-3 could increase the overall sensitivity by only 1.4% [5].

In comparison, high CIN score was found in 77.4% of the recurrent cases while not detected in healthy controls, which was significantly higher than that of CA15-3 and CEA. Notably, the combinations of cfDNA CIN with two biomarkers further increased the positive detection rate to 88.7%. Furthermore, the correlation analysis demonstrated that CIN score was not associated with the expression level of CA15-3/CEA, indicating it to be independent of two biomarkers. The above results suggested that CIN score might be used as a non-invasive peripheral biomarker to monitor

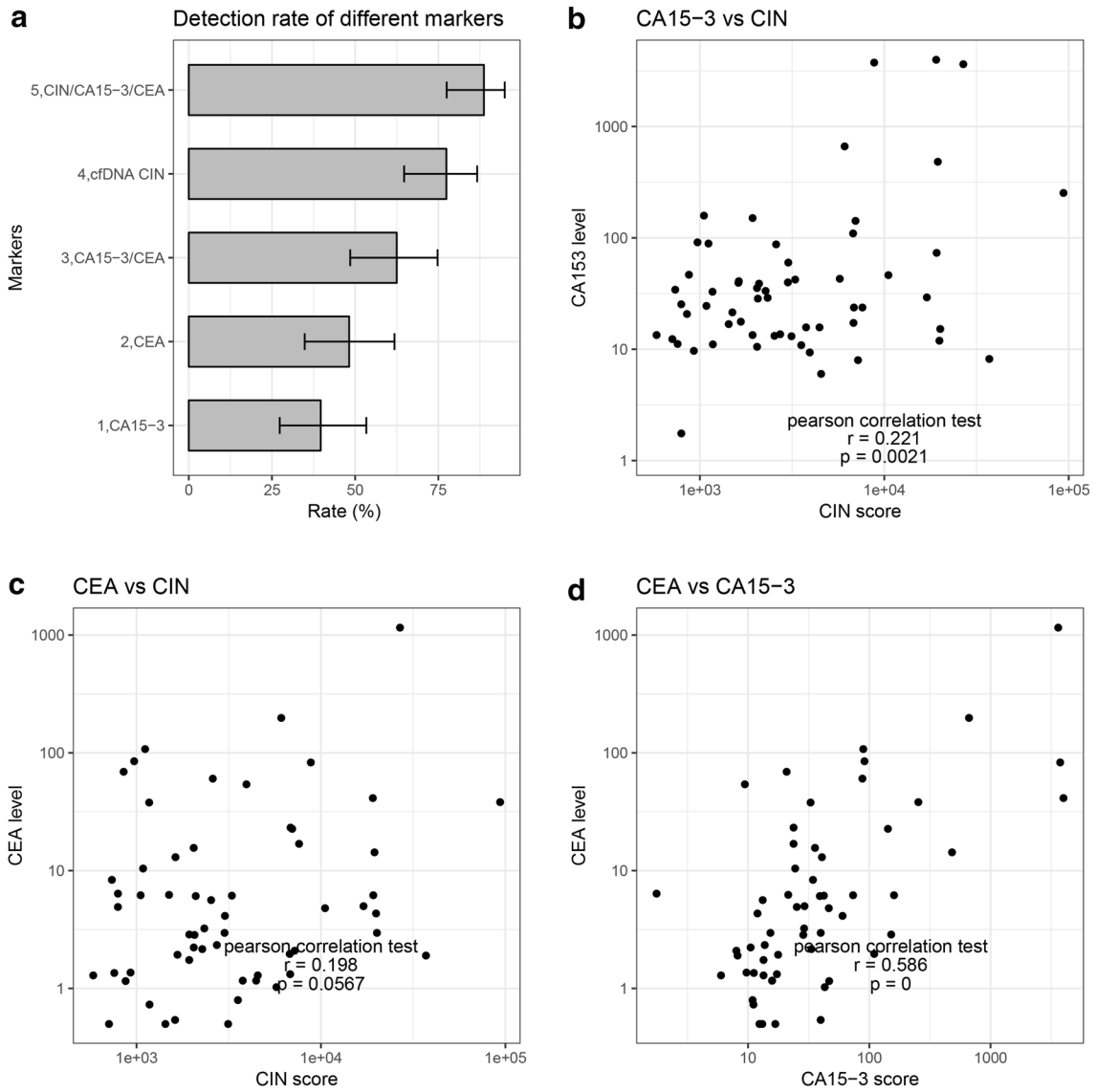

Fig. 2 **a** Detection results of CA15-3, CEA, cfDNA chromosomal instability (CIN) and combination to monitor breast cancer recurrence. **b–d** Correlations between CA15-3 level and CIN score, CEA level and CIN score, and CA15-3 level and CEA level, respectively

Table 3 Detection rate of cfDNA CIN for recurrent breast cancer patients with different levels of CA15-3 and CEA

	CA15-3 ≥ 35 CEA ≥ 5	CA15-3 ≥ 35 CEA < 5	CA15-3 < 35 CEA ≥ 5	CA15-3 < 35 CEA < 5	CA15-3, CEA Unknown	All pts
	N = 15	*N* = 8	*N* = 11	*N* = 21	*N* = 3	*N* = 58
cfDNA CIN positive	12	7	6	17	3	45
cfDNA CIN negative	3	1	5	4	0	13
Rate %	80.0	87.5	54.5%	81.0%	100%	77.6%
(95% CI)	(67.6–100%)	(51.0–100%)	(15.0–85.0%)	(57.8–95.7%)	(20.8–93.6%)	(73.4–95.3%)

cfDNA cell-free DNA, *CIN* chromosomal instability, *CI* confidence interval

Table 4 Levels of CA15-3, CEA, and cfDNA CIN in the recurrent breast cancer patients with visceral or non-visceral metastasis

	Visceral metastasis ($n=45$)	Non-visceral metastasis ($n=16$)	Unknown ($n=1$)	Total ($n=62$)	P value
UCAD cfDNA CIN high	34	13	1	48	Not significant[a]
UCAD cfDNA CIN low	11	3	0	14	
Average CIN score	6836	7522	/	7198	
Rate (95% CI)	75.6 (71.1–97.4)%	81.3 (55.2–95.3)%	/	77.4 (73.4–95.3)%	
CA15-3 ≥ 35 (KU/L)	20	3	0	12	0.03[b]
CA15-3 < 35 (KU/L)	25	13	0	19	
Average CA15-3 (KU/L)	245.6	26.0	/	184.7	
Rate (95% CI)	44.4 (27.3–68.3)%	18.8 (4.70–44.8)%	/	38.7 (23.7–56.2)%	
CEA ≥ 5 (ng/ml)	19	8	0	13	Not significant[c]
CEA < 5 (ng/ml)	26	8	0	18	
Average CEA (ng/ml)	46.3	17.9	/	38.8	
Rate (95% CI)	42.2 (30.0–80.7)%	50 (25.4–74.6)%	/	41.9 (26.4–59.2)%	

UCAD ultrasensitive chromosomal aneuploidy detector, *cfDNA* cell-free DNA, *CIN* chromosomal instability, *CI* confidence interval, *NS* no significance

[a]Difference in the CIN score between the recurrent breast cancer patients with visceral or non-visceral metastasis

[b]Difference in the CA15-3 level between the recurrent breast cancer patients with visceral or non-visceral metastasis

[c]Difference in the CEA level between the recurrent breast cancer patients with visceral or non-visceral metastasis

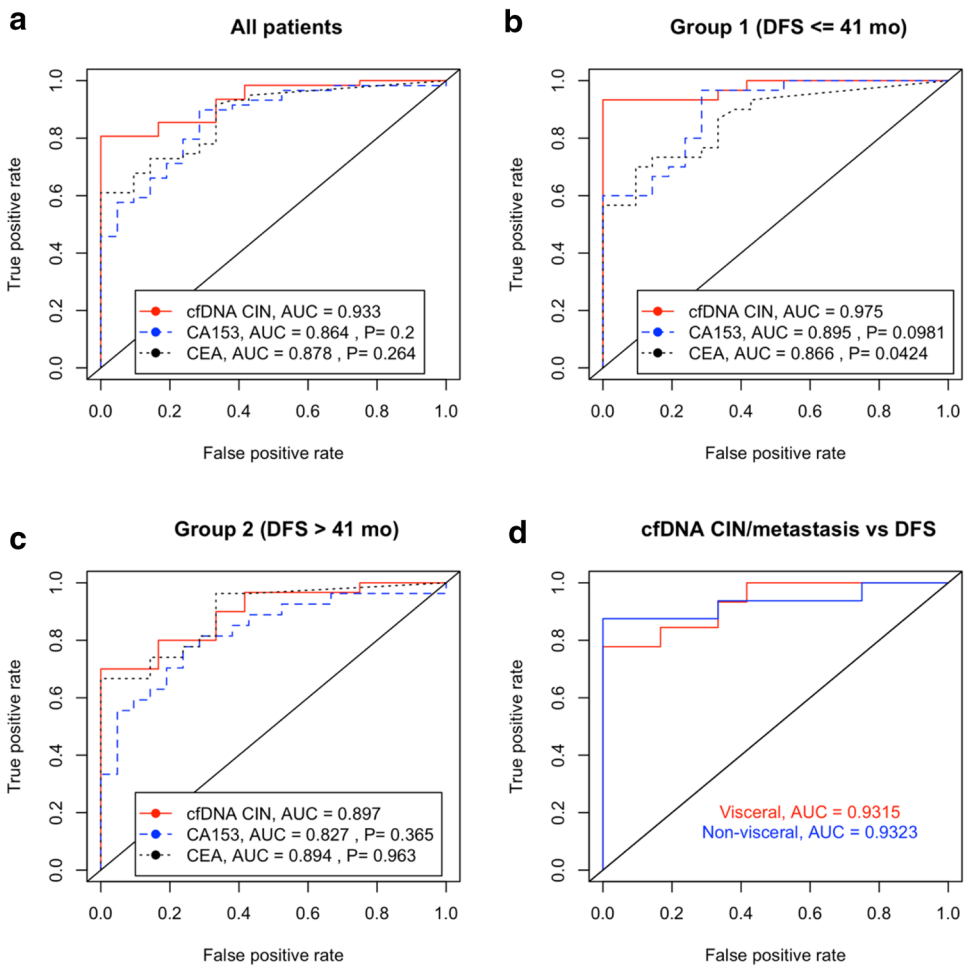

Fig. 3 Performance of each biomarker to predict breast cancer relapses for all patients, patients with DFS ≤ 41 months and patients with DFS > 41 months, respectively. **a** cfDNA chromosomal instability (CIN). **b** CA15-3. **c** CEA. **d** cfDNA CIN to predict breast cancer patients with visceral or non-visceral metastasis

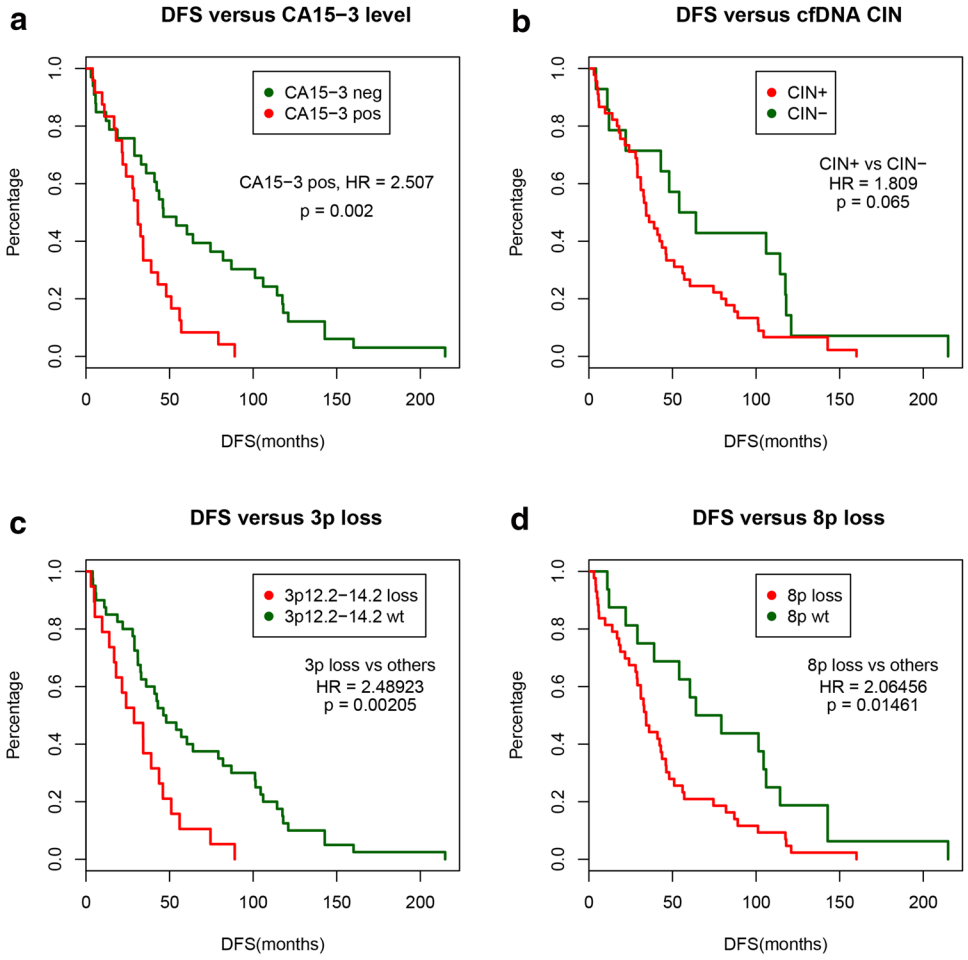

Fig. 4 Correlations between DFS and **a** CA15-3 level (positive/negative). **b** Plasma cfDNA chromosomal instability (CIN, high/low). **c** Copy number variation on chromosome 3p (loss/wildtype). **d** Copy number variation on chromosome 8p (loss/wildtype)

breast cancer recurrence competent to the traditional biomarkers (CA15-3 and CEA).

Sandri et al. [28] reported that pre-surgical level of CA15-3 was used as an independent prognostic factor for metastases and deaths, which was recommend in the adjuvant therapeutic algorithm. In our study, shorter DFS was associated with higher CA15-3 level, which revealed its value in monitoring the relapse of breast cancer. Also, shorter DFS tended to correlate with higher CIN score though no statistically significance was observed. Just four patients were CIN-low in our cohort, and the small number of patients led to no significance in statistics. Similarly, Stover et al. [13] found that cfDNA tumor fraction was related to worse metastatic survival among triple-negative breast cancer patients. In addition, Marcel et al. discovered the CIN score was significantly associated with prognosis in ER-positive, luminal B, and her2/neu subtypes, but not in ER-negative patients [29]. The majority of molecular types were luminal A and luminal B in our cohort, which accounting for 87.1% (54/62) of total patients. The value of CIN score with cancer related survival in non-triple-negative breast cancer patient needs to be further validated.

Furthermore, we found that copy losses on Chr 3p and Chr 8p were significantly associated with DFS. In our study, Chr 1q amplification was also frequently found among the recurrent breast cancer patients, with the frequency of 58.1%. Similarly, Goh et al. [12] identified that Chr 1q21.3 amplification detected from cfDNA was strongly associated with the early relapse in breast cancer patients, with the mechanism that 1q21.3-directed S100A7/8/9-IRAK1 feedback loop drives the tumor growth. Further study should be conducted to investigate the potential mechanism of Chr3p loss and Chr8p loss shortening DFS.

In this study, we also detected several gene amplifications among the recurrent breast cancer patients, which may be useful to guide the clinical practices. Recent work implicated ESR1 mutations was detected in 34% patients with ER + metastatic breast cancer and tumor cells with ESR1 loss may less dependent of estrogen receptor signaling, which can lead to hormone resistance [30]. For instance, ESR1 copy loss was frequently found in patients whose primary tumor with ER-positive. Another well-studied gene, FGFR1, was found to be amplified in 17.6% of the patients, which has been proven to play an important role in breast

cancer development [31]. Similarly, Turner N found FGFR1 amplification was found in 16% to 27% of luminal B-type breast cancers [32]. Moreover, other genes such as OXR1, IKBKB, MDM2, CCND1, PCDHA were found with amplification in 24.2%, 25.8%, 25.8%, 26.5%, and 29% of the patients, respectively. Amplification of CCND1 was found in 21% ER + metastatic breast cancer patients [30]. IKBKE was identified as a human breast cancer oncogene and amplified and overexpressed in a considerable proportion of human breast tumors [33]. However, the roles of these genes in breast cancer have not been well understood, for which functional validation remains to be conducted in the future.

In summary, plasma cfDNA CIN detection revealed better performance in monitoring the breast cancer relapse, superior to the traditional biomarkers (CA15-3 and CEA), especially for biomarker-negative patients. The combination of cfDNA CIN with traditional biomarkers further improved the detection rate of recurrent breast cancer patients, which may provide a new method for monitoring the early relapse. Moreover, high cfDNA CIN tended to be found in patients with shorter DFS which may provide potential therapeutical targets for patients. Yet, further investigations with large sample size are highly warranted.

Acknowledgements We would like to express special thanks to all participating patients and their families.

Funding This trial was funded by grants from National Natural Science Foundation of China (Grant Number: 81672597), Natural Science Foundation of Zhejiang Province, China (Grant Number: LY17H160038, LY14H160030), Key Research-Development Program of Zhejiang Province [Grant Number: 2019C04001, 2017C03013], Science and Technology Program offered by the Health Bureau of Zhejiang Province, China (Grant Number: 2017RC014), and Joint Key Program of Zhejiang Province-Ministry of Health [Grant Number: WKJ-ZJ-1714], and Qianjiang Talents Fund of Zhejiang Province [Grant Number: QJD1602026].

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the Clinical Research Ethics Committee of Zhejiang Cancer Hospital. All experiments comply with the current China laws.

Informed Consent Written informed consent was obtained from all individual participants included in the study.

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. CA 68(6):394–424
2. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray F (2015) Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer 136(5):E359–E386. https://doi.org/10.1002/ijc.29210
3. Lacey H, Judith H (2015) MUC1-mediated motility in breast cancer: a review highlighting the role of the MUC1/ICAM-1/Src signaling triad. Clin Exp Metastasis 32(4):393–403
4. Dawson SJ, Tsui DW, Murtaza M, Biggs H (2013) Analysis of circulating tumor DNA to monitor metastatic breast cancer. N Engl J Med 368(13):1199–1209
5. Guadagni F, Ferroni P, Carlini S, Mariotti S, Spila A, Aloe S, D'Alessandro R, Carone MD, Cicchetti A, Ricciotti A (2001) A re-evaluation of carcinoembryonic antigen (CEA) as a serum marker for breast cancer: a prospective longitudinal study. Clin Cancer Res 7(8):2357–2362
6. Aguiar-Bujanda D, Bohn-Sarmiento U, Aguiar-Morales J (2004) False elevation of serum CA 15-3 levels in patients under follow-up for breast cancer. Breast J 10(4):375–376
7. Gioia D, Di Stieber P, Schmidt GP, Nagel D, Heinemann V, Baur-Melnyk A (2015) Early detection of metastatic disease in asymptomatic breast cancer patients with whole-body imaging and defined tumour marker increase. Breast Dis 112(5):809–818
8. Newman AM, Bratman SV, To J, Wynne JF, Eclov NC, Modlin LA, Liu CL, Neal JW, Wakelee HA, Merritt RE, Shrager JB, Loo BW Jr (2014) An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage. Nat Med 20(5):548–554. https://doi.org/10.1038/nm.3519
9. Yarchoan M, Hopkins A, Jaffee EM (2017) Tumor mutational burden and response rate to PD-1 inhibition. N Engl J Med 377(25):2500
10. Network TCGA (2012) Comprehensive molecular portraits of human breast tumors. Nature 490(7418):61–70
11. Knouse KA, Davoli T, Elledge SJ, Amon A (2017) Aneuploidy in cancer: seq-ing answers to old questions. Annu Rev Cancer Biol 1(1):335–354
12. Goh JY, Feng M, Wang W, Oguz G, Smjm Y, Lee PL, Bao Y, Lim TH, Wang P, Tam WL (2017) Chromosome 1q21.3 amplification is a trackable biomarker and actionable target for breast cancer recurrence. Nat Med 23(11):1319
13. Stover DG, Parsons HA, Ha G, Freeman SS, Barry WT, Guo H, Choudhury AD, Gydush G, Reed SC, Rhoades J (2018) Association of cell-free DNA tumor fraction and somatic copy number alterations with survival in metastatic triple-negative breast cancer. J Clin Oncol 36(6):543–553
14. Peiyong J, Chan CWM, Chan KCA, Suk Hang C, John W, Vincent Wai-Sun W, Wong GLH, Chan SL, Mok TSK, Chan HLY (2015) Lengthening and shortening of plasma DNA in hepatocellular carcinoma patients. Proc Natl Acad Sci USA 112(11):E1317
15. Adalsteinsson VA, Ha G, Freeman SS, Choudhury AD, Stover DG, Parsons HA, Gydush G, Reed SC, Rotem D, Rhoades J (2017) Scalable whole-exome sequencing of cell-free DNA reveals high concordance with metastatic tumors. Nat Commun 8(1):1324
16. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, Müller M (2011) pROC: an open-source package for R and S + to analyze and compare ROC curves. BMC Bioinform 12:1–8
17. Morenobetancur M, Sadaoui H, Piffaretti C, Rey G (2017) Survival analysis with multiple causes of death. Epidemiology 28(1):12
18. Lange AM, Lo HW (2018) Inhibiting TRK proteins in clinical cancer therapy. Cancers 10(4):105
19. Razavi P, Chang MT, Xu G, Bandlamudi C, Ross DS, Vasan N, Cai Y, Bielski CM, Donoghue MTA, Jonsson P, Penson A, Shen

R, Pareja F, Kundra R, Middha S, Cheng ML, Zehir A, Kandoth C, Patel R, Huberman K, Smyth LM, Jhaveri K, Modi S, Traina TA, Dang C, Zhang W, Weigelt B, Li BT, Ladanyi M, Hyman DM, Schultz N, Robson ME, Hudis C, Brogi E, Viale A, Norton L, Dickler MN, Berger MF, Iacobuzio-Donahue CA, Chandarlapaty S, Scaltriti M, Reis-Filho JS, Solit DB, Taylor BS, Baselga J (2018) The genomic landscape of endocrine-resistant advanced breast cancers. Cancer Cell 34(3):427–438. https://doi.org/10.1016/j.ccell.2018.08.008

20. Kjaer IM, Bechmann T, Brandslund I, Madsen JS (2017) Prognostic and predictive value of EGFR and EGFR-ligands in blood of breast cancer patients: a systematic review. Clin Chem Lab Med 56(5):688–701
21. Beardsley DI, Kowbel D, Lataxes TA, Mannino JM, Xin H, Kim WJ, Collins C, Brown KD (2003) Characterization of the novel amplified in breast cancer-1 (NABC1) gene product. Exp Cell Res 290:402–413
22. Ramaiah MJ, Vaishnave S (2018) BMI1 and PTEN are key determinants of breast cancer therapy: a plausible therapeutic target in breast cancer. Gene 678:302–311
23. Rivand M, Khorrami MS, Fiuji H, Shahidsales S, Hasanzadeh M, Jazayeri MH, Hassanian SM, Ferns GA, Saghafi N, Avan A (2017) The 9p21 locus: a potential therapeutic target and prognostic marker in breast cancer. J Cell Physiol 233(7):5170–5179
24. Johnson J, Thijssen B, Mcdermott U, Garnett M, Wessels LFA, Bernards R (2016) Targeting the RB-E2F pathway in breast cancer. Oncogene 35(37):4829–4835
25. Schon K, Tischkowitz M (2018) Clinical implications of germline mutations in breast cancer: TP53. Breast Cancer Res Treat 167(2):417–423
26. Chen WZ, Shen JF, Zhou Y, Chen XY, Liu JM, Liu ZL (2017) Clinical characteristics and risk factors for developing bone metastases in patients with breast cancer. Sci Rep 7(1):11325
27. Li X, Dai D, Chen B, Tang H, Xie X, Wei W (2018) Clinicopathological and prognostic significance of cancer antigen 15-3 and carcinoembryonic antigen in breast cancer: a meta-analysis including 12,993 patients. Dis Markers 2018:1–15
28. Sandri MT, Salvatici M, Botteri E, Passerini R, Zorzino L, Rotmensz N (2012) Prognostic role of CA15.3 in 7942 patients with operable breast cancer. Breast Cancer Res Treat 132(1):317–326
29. Marcel S, Marlous H, Sieuwerts AM, Stefan S, Yi Z, Yixin W, Foekens JA, Martens JWM (2011) Patterns and incidence of chromosomal instability and their prognostic relevance in breast cancer subtypes. Breast Cancer Res Treat 128(1):23–30
30. Toy W, Shen Y, Won H, Green B, Sakr RA, Will M, Li Z, Gala K, Fanning S, King TA (2013) ESR1 ligand-binding domain mutations in hormone-resistant breast cancer. Nat Genet 45(12):1439
31. Babina IS, Turner NC (2017) Advances and challenges in targeting FGFR signalling in cancer. Nat Rev Cancer 17(5):318
32. Nicholas T, Alex P, Rachel S, Maryou L, Felipe G, Lopez-Garcia MA, Rachael N, Caterina M, Elizabeth I, Alan M (2010) FGFR1 amplification drives endocrine therapy resistance and is a therapeutic target in breast cancer. Cancer Res 70(5):2085
33. Boehm JS, Zhao JJ, Jun Y, So Young K, Ron F, Dunn IF, Sjostrom SK, Garraway LA, Stanislawa W, Richardson AL (2007) Integrative genomic approaches identify IKBKE as a breast cancer oncogene. Cell 129(6):1065–1079