Detection of ctDNA with Personalized Molecular Barcode NGS and Its Clinical Significance in Patients with Early Breast Cancer

Tetsuhiro Yoshinami a , Naofumi Kagara a,⁎, Daisuke Motooka b , Shota Nakamura b , Tomohiro Miyake a , Tomonori Tanei a , Yasuto Naoi a , Masafumi Shimoda a , Kenzo Shimazu a , Seung Jin Kim a , Shinzaburo Noguchi a,c

a Department of Breast and Endocrine Surgery, Osaka University Graduate School of Medicine, 2-2-E10 Yamadaoka, Suita, Osaka 565-0871, Japan
b Genome Information Research Center, Research Institute for Microbial Diseases, Osaka University, 3-1 Yamadaoka, Suita, Osaka 565-0871, Japan
c Department of Breast and Endocrine Surgery, Osaka University Graduate School of Medicine, 2-2-E10 Yamadaoka, Suita, Osaka 565-0871, Japan

ABSTRACT

We attempted to detect circulating tumor DNA (ctDNA), taking advantage of molecular barcode next-generation sequencing (MB-NGS), which can be more easily customized to detect a variety of mutations with a high sensitivity than PCR-based methods. Sequencing with a gene panel consisting of the 13 most frequently mutated genes in breast tumors from stage I or II patients revealed 95 somatic mutations in the 12 genes in 62% (62/100) of tumors. Then, plasma DNA from each patient (n = 62) before surgery was analyzed via MB-NGS customized to each somatic mutation, resulting in the detection of ctDNA in 16.1% (10/62) of patients. ctDNA was significantly associated with biologically aggressive phenotypes, including large tumor size (P = .004), positive lymph node (P = .009), high histological grade (P < .001), negative ER (P = .018), negative PR (P = .017), and positive HER2 (P = .046). Furthermore, distant disease-free survival was significantly worse in patients with ctDNA (n = 10) than those without ctDNA (n = 52) (P < .001). Our results demonstrate that MB-NGS personalized to each mutation can detect ctDNA with a high sensitivity in early breast cancer patients at diagnosis, and it seems to have a potential to serve as a clinically useful tumor marker for predicting their prognosis.

Introduction

Circulating tumor DNA (ctDNA) is a DNA fragment released from tumor cells into blood, and it harbors genomic alterations specific to original tumor cells [1–3]. The clinical usefulness of ctDNA in breast cancer patients has been widely studied in the fields of diagnosis, prognosis prediction, and monitoring of recurrence or therapeutic effect, and its superiority over conventional serum tumor markers has been reported [4–8].

To detect ctDNA, a highly sensitive approach is required due to its very low amount in blood [2,3]. Moreover, because breast cancer is a heterogeneous disease harboring various alterations [9], methods for ctDNA detection must be available for any type of mutations. Currently, digital PCR (dPCR) has been widely used for ctDNA detection due to its high sensitivity and technical ease [10–12]. However, it is only useful for predetermined mutations and not appropriate for the random mutations [13]. On the other hand, next-generation sequencing (NGS) can detect unknown mutations, and the design and optimization for target sequencing is simple. Although conventional NGS has a limited detection sensitivity [3], recent molecular barcode (MB) technology has enabled NGS to detect ctDNA with a sufficient sensitivity at a variant allele frequency (VAF) of approximately 0.1% [14–16]. Thus, molecular barcode NGS (MB-NGS) is considered suitable for the detection of rare ctDNA in patients with early breast cancers by targeting their individual mutations.

In the present study, we screened the mutations in 100 early breast cancer patients with a panel sequencing and aimed to detect ctDNA by MB-NGS targeting tumor-specific mutations in plasma. The ctDNA status was analyzed with reference to the clinicopathological characteristics of the patients, and its prognostic or diagnostic significance was investigated.

Materials and Methods

Patients and Samples

One hundred patients with stage I or II breast cancer who received surgery without preoperative systemic therapies at Osaka University Hospital between 2007 and 2012 were included in the present study. The clinicopathological characteristics of the patients are shown in Table 1. Fresh frozen tissues were obtained at the time of surgery, and peripheral blood
Table 1  
Clinicopathological features of the breast tumors included in the present study

|                | Total |
|----------------|-------|
| N              | 100   |
| Age (median)   | 53    |
| pT             | 84    |
| pN             | 73    |
| Stage          | 27    |
| Histological grade | 80 |
| ER Positive    | 79    |
| PGR Positive   | 64    |
| HER2 Positive  | 18    |
| Histology IBC  | 87    |
| Histology ILC  | 5     |
| CEA/CA15-3     | 6     |
| Negative       | 94    |

pT: Pathological tumor size; pN: Pathological lymph node status; ER: Estrogen receptor; PGR: Progesterone receptor; HER2: Human epidermal growth factor receptor 2; IDC: Invasive ductal carcinoma; ILC: Invasive lobular carcinoma.

leukocytes (PBL) and plasma were collected before surgery. Informed consent was obtained before sampling, and this study was approved by the Ethical Review Board of Osaka University Hospital.

**DNA Extraction**

Tumor DNA was extracted from fresh frozen tissue using the DNeasy Blood & Tissue Kit® (Qiagen, Hilden, Germany) according to the manufacturer’s instructions. Plasma was separated from whole blood by centrifugation for 10 min at 3000 rpm (1840 × g) and stored at −80 °C until further use. The samples were centrifuged again for 10 min at 13,300 rpm (16,000 × g) prior to DNA extraction to remove debris. Cell-free DNA (cfDNA) was isolated from 1 mL of plasma using the QIAamp Circulating Nucleic Acid Kit® (Qiagen) according to the manufacturer’s instructions and eluted in 50 μL of AVE buffer (Qiagen). PBL were separated from whole blood by centrifugation for 10 min at 3000 rpm (1840 × g) twice and stored at −80 °C until further use in the pellet state. DNA was extracted from the PBL pellet using the DNeasy Blood & Tissue Kit® (Qiagen) according to the manufacturer’s instructions.

**Screening of Mutations in Primary Breast Cancers**

The gene panel was designed by SureDesign (https://earray.chem.agilent.com/suredesign) to cover a whole exon of 13 genes including 12 frequently mutated genes (cBioPortal for cancer genomics, http://www.chiopoortal.org) in breast cancers and ESRI (Supplementary Table 1). The median coverage of sequencing area was 100% (range, 74%-100%) of the whole exon in each gene. Tumor DNA was fragmented by Covaris S220 (Covaris, Inc., Woburn, MA, USA) to 150 bp, and 120 ng of them were used for sequencing. Sequence libraries were prepared with a custom SureSelect XT HS Target Enrichment System (Agilent Technologies, Inc. Santa Clara, CA, USA) according to the manufacturer’s instructions and sequenced with Illumina MiSeq (Illumina, San Diego, CA, USA). SureCall ver4.0 (https://www.agilent.com/en/download-software-surecall) was used for variant calling. The exclusion criteria were as follows; in introns, with VAF of <5%, with a depth of <300 reads, or reported in dbSNP/1000G database with a population of ≥1%. Variants obtained by panel sequencing were confirmed by Sanger sequencing using paired PBL DNA to exclude the patient-specific SNPs. Four hotspot mutations (AKT1-E17K and PIK3CA-E542K/E545K/H1047R) with VAF ≥1% were also included.

**Detection of cfDNA in Plasma with Personalized MB-NGS**

The detection of cfDNA in plasma was performed with MB-NGS targeting the specific mutations detected in the primary tumor of each patient. If two or more mutations were found in the primary tumors, the most frequent and the second most frequent mutations were analyzed by MB-NGS when sufficient samples were available. To detect cfDNA, two libraries of each target mutation were separately prepared using plasma DNA and analyzed using MiSeq or HiSeq (Illumina) as previously described [14]. In brief, an assignment of 15-base (BDHVBDBHVBDBDBH) MB and adaptors (Rd1SP and Rd2SP) was performed at the first PCR with personalized primers. The primer sequences are shown in Supplementary table 2. PCR was performed in a 40-μL reaction containing 20 μL of template DNA, 5× Phusion HF buffer (NEB, Ipswich, MA, USA), 0.9 units of Phusion polymerase (NEB), 250 μM dNTPs, and 0.01 μM of each primer. The cycling conditions were one cycle of 98 °C for 30 s; 15 cycles of 98 °C for 10 s, 57 or 63 °C for 2 min, and 72 °C for 10 s; and one cycle of 72 °C for 10 min. To remove the first PCR primers, each product was digested with 25 units of Exonuclease-I and 10× Exonuclease-I reaction buffer (NEB) in a 50-μL reaction at 37 °C for 1 h and 5 min heat inactivation at 98 °C. Adapter primers with P5 and P7 sequences were added at 0.01 μM each to the first PCR product, and the second PCR was performed in a 56-μL reaction. The cycling conditions were one cycle of 98 °C for 30 s; 10 cycles of 98 °C for 10 s, 55 °C for 30 s, and 72 °C for 30 s; and one cycle of 72 °C for 10 min. The second PCR products were purified using AMPure XP (Beckman Coulter, Brea, CA, USA) according to the manufacturer’s instructions and eluted in 10 μL of nuclease-free water (Qiagen). The concentration of Ampure XP was adjusted depending on amplicon sizes ranging from 0.8 × to 1.4 ×. The third PCR was performed using P5/P7 primers (0.5 μM each) in a 20-μL reaction containing 10-μL of the second PCR product, 5× Phusion HF buffer (NEB), 0.9 units of Phusion polymerase (NEB), and 250 μM dNTPs. The cycling conditions were one cycle of 98 °C for 30 s; 30 cycles of 98 °C for 10 s, 60 °C for 30 s, and 72 °C for 30 s; and one cycle of 72 °C for 10 min. The third PCR products were purified twice using AMPure XP (Beckman Coulter) and eluted in 15 μL of nuclease-free water (Qiagen). Libraries were quantified using a Labchip GX Touch instrument (Waltham, MA, USA, Perkin Elmer) or Agilent 2100 Bioanalyzer instrument and High Sensitivity DNA Assay (Agilent Technologies).

**Data Analysis for Single Nucleotide Variant and Insertion/Deletion**

The variant detection analysis for a single nucleotide variant (SNV) was performed in a similar manner to that described in our previous study [14]. Each MB family including over 30 reads was analyzed. For insertion/deletion (In/Del) analysis, the assembled paired-end reads with unideal lengths (shorter or longer than the designed amplicon) were extracted using PEAR 0.9.6. MB sequences were removed from the extracted assembled sequences and then clustered into each MB family containing sequences with the same MB sequence by CD-HIT-EST and the custom Ruby script. For each MB family including over 30 reads, sequences were mapped onto each reference sequence using BWA ver.0.7.5. The VAF at each variant was calculated by dividing the number of MB families with the variant (variant MB families) by the number of total MB families. In/Del analysis was performed using lofreq2 [17]. When SNV was positive in two, one, or no libraries, the samples were classified as “double-positive,” “single-positive,” or “double-negative,” respectively. Only “double-positive” samples with VAF >0.1% were judged as ctDNA positive, and all others were judged as ctDNA negative. In the case of In/Dels, they were analyzed without MB, and ctDNA was judged as positive when at least one library included mutant alleles with VAF >0.1%. To rule out the clonal hematopoiesis in ctDNA-positive patients, personalized MB-NGS was performed using DNA of the paired PBL.
Results

Identification of Individual Mutations in Primary Breast Cancers

Whole exon sequencing with 13 gene panels was performed in 100 primary breast cancers, and 95 mutations were detected in 62 (62%) tumors; one mutation was detected in 37 (37%) tumors, two mutations were detected in 20 (20%) tumors, three mutations were detected in four (4%) tumors, and six mutations were detected in one (1%) tumor (Figure 1A). The median read depth was 1108 (range; 719-2098), and the median mutant allele frequencies (AF) was 27.9% (range; 1.5%-68.5%). PIK3CA (32%) and TP53 (18%) were the most frequent mutations, followed by GATA3 (11%), AKTI (8%), MAP3K1 (5%), and CDH1 (4%) (Figure 1B). The most frequent mutations by subtype were PIK3CA (37%) and GATA3 (13%) in estrogen receptor (ER)- or progesterone receptor (PgR)-positive / human epidermal growth factor receptor 2 overexpression (HER2)-negative tumors, PIK3CA (28%) and TP53 (28%) in HER2-positive tumors, and TP53 (60%) in ER-negative / PgR-negative / HER2-negative tumors (Supplementary Figure 1). The 95 mutations detected comprised 35 different types of SNVs (n = 69) and 22 types of In/Dels (n = 26), and their distribution is presented in Figure 2. Multiple mutations were detected in several hotspots; 15 mutations in PIK3CA-H1047R, eight in PIK3CA-E545K and AKTI-E17K, five in PIK3CA-E542K and GATA3-P409fs, and two in TP53-Y107* and -R248Q. The remaining 50 mutations were found only once at each site.

Sensitivity of ctDNA Detection by MB-NGS Targeting Individual Mutations in Primary Tumors

Forty-one sets of primers for MB-NGS were designed to cover 95 mutations (Figure 2). Background errors were analyzed for the MB-NGS primers using 3 ng of plasma DNA from a healthy individual. In the analysis of 24 amplicons used to detect SNVs in plasma, the median number of MB families of each amplicon was 5261 (range, 1141-13,579), which was sufficient to obtain a detection sensitivity of 0.1%. In conventional analyses without MB, the background errors were observed in all 24 target SNVs with a VAF ranging from 0.001 to 0.187%. After the MB analyses, background SNVs were completely removed in 23 SNVs except TP53-R306* with a frequency of 0.076% (Figure 3). In the analysis of In/Dels, the median read depth of each amplicon was 931,137 (range, 615,910 – 1,110,008), and the background errors were not detected at all in 10 In/Dels even without MB analyses. Based on these results, SNVs and In/Dels were analyzed with and without MB, respectively, and the detection limit was set at 0.1% for both in the following analyses.

Detection of ctDNA by MB-NGS in Early Breast Cancer Patients

Plasma DNA from 62 patients identified as having mutations in breast cancer was analyzed to search for ctDNA with patient-specific mutations. Of the 62 samples, 54 were subjected to MB-NGS for a single mutation and eight for double mutations. The median amount of input plasma DNA per library was 4.9 ng (range, 2.1–13.2 ng). For SNVs, 10 SNVs from nine plasma samples were detected as “double-positive” with VAF >0.1% (range, 0.10% - 2.51%), with all being judged as ctDNA-positive. Eleven SNVs from nine plasma samples were detected as “single-positive,” and 37 SNVs from 33 plasma samples were “double-negative,” with all being judged as ctDNA-negative. For In/Dels, one deletion (TP53-D208fs, VAF = 0.20%) was detected in one plasma sample. In total, 11 mutations were detected in 10 of 62 plasma samples (16.1%); two SNVs in one plasma sample, one SNV in eight plasma samples, and one In/Del in one plasma sample (Figure 4 and Table 2). The DNA from the paired PBL of these 10 patients was analyzed by MB-NGS, and no clonal hematopoiesis was observed.

Clinical Significance of ctDNA in Early Breast Cancer Patients

Clinical and pathological factors were compared between ctDNA-positive (n = 10) and ctDNA-negative (n = 52) patients (Table 3). ctDNA was significantly more frequent in the patients whose tumor was larger in size (P < .004) or that were node-positive (P = .009), higher stage (P < .001), higher in histological grade (P < .001), ER-negative (P = .018), PgR-negative (P = .017), and HER2-positive (P = .046). Histology and serum tumor markers such as CEA/CA15–3 were not associated with ctDNA status. ctDNA was significantly more positive in stage II breast cancer patients than serum tumor markers (39% vs 4%, P = .013) but not in stage I breast cancer patients (3% vs 10%, P = .371) (Supplementary Figure 2). Seven of the 62 patients experienced distant recurrence, and ctDNA was positive in four of them. The prognostic curve analysis showed that distant disease-free survival was significantly worse in ctDNA-positive patients (n = 10) than in ctDNA-negative patients (n = 52) (P < .001) (Figure 5). Positive ctDNA was significantly associated with worse distant disease-free survival compared to patients with ctDNA-negative status.

Statistics

R 3.4.1 was used for statistical processing. Fisher’s exact test was used to compare 2 x 2 groups to apply the 2 x 2 contingency tables with a dichotomous trait. The Kaplan–Meier approach was performed to estimate the survival function, and a log-rank test was used to compare the survival distributions of the two groups. A univariate analysis of association between survival trait. The Kaplan–Meier approach was performed to estimate the survival function, and a log-rank test was used to compare the survival distributions of the two groups. A univariate analysis of association between survival need to be conducted with the Cox proportional hazard model. P < .05 was considered significant.

Figure 1. Mutations detected in 100 primary breast tumors. The proportion of the number of mutations detected in 100 primary breast tumors by 13-gene panel sequencing (A) and the frequencies of mutations of each gene (B) are presented.

Frequency

| Gene     | 1 gene | 2 genes | 3 genes | 0 gene |
|----------|--------|---------|---------|--------|
| TP53     | 18%    | 11%     | 8%      | 5%     |
| GATA3    | 5%     | 4%      | 3%      | 2%     |
| AKTI     | 2%     | 2%      | 1%      | 1%     |
| MAP3K1   | 1%     | 0%      | 0%      | 0%     |
| CDH1     | 1%     | 0%      | 0%      | 0%     |
| PIK3CA   | 32%    | 30%     | 28%     | 25%    |
| TP53     | 11%    | 10%     | 9%      | 8%     |
| GATA3    | 5%     | 4%      | 3%      | 2%     |
| AKTI     | 2%     | 2%      | 1%      | 1%     |
| MAP3K1   | 1%     | 0%      | 0%      | 0%     |
| CDH1     | 1%     | 0%      | 0%      | 0%     |
| PIK3CA   | 32%    | 30%     | 28%     | 25%    |
| TP53     | 11%    | 10%     | 9%      | 8%     |
| GATA3    | 5%     | 4%      | 3%      | 2%     |
| AKTI     | 2%     | 2%      | 1%      | 1%     |
| MAP3K1   | 1%     | 0%      | 0%      | 0%     |
| CDH1     | 1%     | 0%      | 0%      | 0%     |
Figure 2. Distribution of 95 somatic mutations and the locations of personalized MB-NGS primers. Fifty-seven types of 95 somatic mutations were detected in 62 primary breast tumors, and their locations are plotted in the lollipop figure. Y-axis indicated the number of mutations. Black arrows indicate the locations of 40 personalized MB-NGS primers designed to cover all somatic mutations. MB-NGS, molecular barcode-next generation sequencing.
Figure 3. Background error frequencies in a healthy individual plasma at each target single nucleotide variant with or without molecular barcode analysis.

Background error frequencies at the 24 target mutations obtained by conventional (NGS) (white dots) and MB-NGS (black dots) using cell-free DNA from a healthy control plasma are plotted in descending order of frequency by conventional NGS. MB, molecular barcode. NGS, next generation sequencing. MB-NGS, molecular barcode-next generation sequencing.

Figure 4. Variant allele frequency of target mutations in cell-free DNA and paired PBL DNA. The variant allele frequency of target mutations detected in plasma DNA (black dots) and the paired PBL DNA (white dots) of ctDNA-positive patients by MB-NGS are plotted. PBL, peripheral blood leukocytes. ctDNA, circulating tumor DNA. MB-NGS, molecular barcode-next generation sequencing.

Table 2

| Case | Gene | AA change | Tissue VAF | Plasma VAF | Variant MB families | Total MB families | Clonal hematopoiesis |
|------|------|-----------|------------|------------|---------------------|------------------|---------------------|
| 10–353 | TP53 | p.R248Q | 49.2% | 2.51% | 330 | 13,150 | Negative |
| 10–353 | TP53 | p.Y107* | 27.8% | 1.75% | 129 | 7387 | Negative |
| 11–202 | PIK3CA | p.H1047R | 38.7% | 0.88% | 184 | 20,818 | Negative |
| 11–187 | TP53 | p.Y107* | 59.9% | 0.86% | 174 | 20,200 | Negative |
| 11–213 | PIK3CA | p.H1047R | 47.8% | 0.85% | 70 | 8250 | Negative |
| 10–486 | TP53 | p.R248Q | 26.9% | 0.51% | 34 | 6613 | Negative |
| 12–566 | PIK3CA | p.H1047R | 42.0% | 0.21% | 13 | 6073 | Negative |
| 11–097 | AKT1 | p.E17K | 62.6% | 0.20% | 69 | 35,326 | Negative |
| 12–407 | TP53 | p.H179Q | 40.6% | 0.10% | 49 | 45,718 | Negative |

MB-NGS, Molecular barcode-next generation sequencing; MB, Molecular barcode; ctDNA, Circulating tumor DNA; AA, Amino acid; VAF, Variant allele frequency.
The distant disease-free survival was compared according to the ctDNA status for 62 patients with mutation-positive breast cancer. Solid line, ctDNA negative (n = 52); dashed line, ctDNA positive (n = 10). ctDNA, circulating tumor DNA.

**Figure 5. ctDNA status and the distant disease-free survival.** The distant disease-free survival was compared according to the ctDNA status for 62 patients with mutation-positive breast cancer. Solid line, ctDNA negative (n = 52); dashed line, ctDNA positive (n = 10). ctDNA, circulating tumor DNA.

**Table 3**

| ctDNA status and clinicopathological features of the breast tumors | ctDNA | P* |
|---|---|---|
| N | 52 | 10 |
| Age (median range) | 51 (36–82) | 55 (35–80) | .886 |
| pT | 1 | 45 | 4 | .004 |
| 2 | 7 | 6 |
| pN | 0 | 43 | 4 | .009 |
| 1 ≤ | 9 | 6 |
| Stage | 1 | 38 | 1 | <.001 |
| II | 14 | 9 |
| Histological grade | 1/2 | 45 | 2 | <.001 |
| 3 | 6 | 8 |
| ER | Unknown | 1 | 0 |
| Positive | 45 | 5 | .018 |
| Negative | 7 | 5 |
| PgR | Positive | 39 | 2 | .017 |
| Negative | 13 | 8 |
| HER2 | Positive | 6 | 4 | .046 |
| Negative | 46 | 6 |
| Histology | IDC | 45 | 10 | 1.000 |
| ILC | 4 | 0 |
| Others | 3 | 0 |
| CEA/CA15–3 | Positive | 4 | 1 | 1.000 |
| Negative | 48 | 9 |

Values are given as number (percentage). ctDNA, Circulating tumor DNA; pT, Pathological tumor size; pN, Pathological lymph node status; ER, Estrogen receptor; PgR, Progesterone receptor; HER2, Human epidermal growth factor receptor 2; IDC, Invasive ductal carcinoma; ILC, Invasive lobular carcinoma.

The analysis of presurgical plasma revealed that ctDNA was positive in 10 (16.1%) of 62 patients and that the clonal hematopoiesis was denied by MB-NGS of the paired PBL DNA. This positive rate of ctDNA was consistent with our previous reports (9.8%-22.7%) measuring ctDNA by targeting MB-NGS of the paired PBL DNA or plasma DNA from a healthy individual (data not shown) and the present study based on the result of control plasma DNA and the amount of input DNA. Recently, several ctDNA assays have been developed with the goal of commercialization [20–22], and our assay exhibited a similar detection sensitivity to them. Given that the background errors on our assay were observed at only one target SNP (TP53 p.R306*) presenting a frequency of 0.076% and were completely suppressed in all of the other SNVs, the cut-off can be set further lower for the majority of mutations by improving the algorithm or by using a greater amount of plasma [23].

The distant disease-free survival by the univariate analysis (P = .004) (Table 4). A multivariate analysis was not conducted due to the small number of events [18].

**Table 4**

| Univariate analysis of various parameters associated with distant disease-free survival | Hazard ratio | 95% CI | P* |
|---|---|---|---|
| pT (2 vs 1) | 0.508 | 0.061–4.229 | .531 |
| pN (1 vs 0) | 5.001 | 1.106–22.69 | .037 |
| Stage (II vs I) | 2.183 | 0.487–9.788 | .308 |
| Histological grade (3 vs 1/2) | 4.542 | 1.009–20.45 | .049 |
| ER (negative vs positive) | 1.597 | 0.307–8.301 | .578 |
| PgR (negative vs positive) | 5.388 | 1.023–28.39 | .047 |
| HER2 status (positive vs negative) | 1.154 | 0.132–10.10 | .897 |
| Adjuvant chemotherapy (Yes vs No) | 1.338 | 0.297–6.400 | .705 |
| CEA/CA15–3 (positive vs negative) | 4.522 | 0.866–23.57 | .073 |
| ctDNA (positive vs negative) | 9.337 | 2.040–42.73 | .004 |

pT, Pathological tumor size; pN, Pathological lymph node status; ER, Estrogen receptor; PgR, Progesterone receptor; HER2, Human epidermal growth factor receptor 2; ctDNA, Circulating tumor DNA.

In the present study, only the patients with “double-positive” results by MB-NGS were considered positive, while there were nine patients with “single-positive” results. Of these nine patients, four were classified into “double-negative” due to the background errors detected in the paired PBL DNA or plasma DNA from a healthy individual (data not shown) and the remaining five were finally classified into “single-positive”. Interestingly, two (40%) of these five “single-positive” patients developed recurrence and, all (n = 47) except one “double-negative” patients were disease-free (Supplementary Figure 3). These results should be interpreted with a great caution due to a very low level of ctDNA and a small number of the patients, but they suggest a possibility that “single positive” patients actually have a very small amount of ctDNA which is associated with poor prognosis.

In conclusion, the current pilot study demonstrated that MB-NGS could detect patient-specific ctDNA targeting diverse mutations in primary breast tumors and indicated that ctDNA can serve as a significant prognostic biomarker. Considering that mutations could not be found in 38% of the patients with our current 13-gene panel, comprehensive analyses such as

**Discussion**

Our study demonstrated that ctDNA can be detected in 16.1% of patients with early breast cancer by means of MB-NGS targeting patient-specific mutations of the primary tumors, and it is associated with the patients' prognosis. Sequencing with the 13-gene panel identified 95 mutations in 62% (62/100) of the primary tumors, which was comparable to the mutation rate ranging from 46% to 59% reported in similar studies using panel sequencing [12,19]. The mutations detected in our study comprised 57 different types of mutations including as many as 50 unrepeated ones, suggesting the diversity of breast cancers.

To address such diverse mutations, MB-NGS was used to detect ctDNA because it is more flexible to design and easier to customize than PCR-based methods. The detection sensitivity was set at 0.1% in the present study based on the result of control plasma DNA and the amount of input DNA. Recently, several ctDNA assays have been developed with the goal of commercialization [20–22], and our assay exhibited a similar detection sensitivity to them. Given that the background errors on our assay were observed at only one target SNP (TP53 p.R306*) presenting a frequency of 0.076% and were completely suppressed in all of the other SNVs, the cut-off can be set further lower for the majority of mutations by improving the algorithm or by using a greater amount of plasma [23].

The analysis of presurgical plasma revealed that ctDNA was positive in 10 (16.1%) of 62 patients and that the clonal hematopoiesis was denied by MB-NGS of the paired PBL DNA. This positive rate of ctDNA was consistent with our previous reports (9.8%-22.7%) measuring ctDNA by targeting methylated DNA using methylation-specific PCR [4,24] or single gene mutation using dPCR [5], indicating a comparable sensitivity of MB-NGS to these PCR-based methods for ctDNA detection. ctDNA was more frequently positive than serum tumor markers (16% vs 6%), especially in stage II breast cancer patients (39% vs 4%, P = .013), suggesting its superiority in diagnosis [25]. Our present data also revealed that the positive ctDNA was significantly associated with worse prognosis, reinforcing the previous findings [4,5,12,26].

In the present study, only the patients with “double-positive” results by MB-NGS were considered positive, while there were nine patients with “single-positive” results. Of these nine patients, four were classified into “double-negative” due to the background errors detected in the paired PBL DNA or plasma DNA from a healthy individual (data not shown) and the remaining five were finally classified into “single-positive”. Interestingly, two (40%) of these five “single-positive” patients developed recurrence and, all (n = 47) except one “double-negative” patients were disease-free (Supplementary Figure 3). These results should be interpreted with a great caution due to a very low level of ctDNA and a small number of the patients, but they suggest a possibility that “single positive” patients actually have a very small amount of ctDNA which is associated with poor prognosis.

In conclusion, the current pilot study demonstrated that MB-NGS could detect patient-specific ctDNA targeting diverse mutations in primary breast tumors and indicated that ctDNA can serve as a significant prognostic biomarker. Considering that mutations could not be found in 38% of the patients with our current 13-gene panel, comprehensive analyses such as
whole exome/genome sequencing will be essential in future study \[10,23,26\] for the application of cDNA assay by MB-NGS to all patients.

Conflicts of interest

Naoi Y. has received honoraria and research funding not related to this study from Sysmex and holds the joint patents not related to this study with Shimazu and holds the joint patents not related to this study with Sysmex. Shimazu K. has received honoraria not related to this study from Sysmex and holds the joint patents not related to this study with Sysmex. No potential conflicts of interest were disclosed by the other authors.

Supplementary data to this article can be found online at doi.org/10.1111/j.1365-2303.2020.150787.

Acknowledgments

The authors wish to thank all the patients who took part in this study.

Funding

This work was supported by JSPS KAKENHI Grant Number JP19K09069.

References

[1] F. Diehl, K. Schmidt, M.A. Choti, K. Romans, S. Goodman, M. Li, K. Thornton, N. Agrawal, L. Sokol, S.A. Szabo, K.W. Kinzler, B. Vogelstein, L.A. Diaz Jr., Circulating mutant DNA to assess tumor dynamics, Nat. Med. 14 (2008) 985–990.

[2] L.A. Diaz Jr., A. Bardelli, Liquid biopsies: genotyping circulating tumor DNA, J. Clin. Oncol. 32 (2014) 579–586.

[3] R.B. Corcoran, B.A. Chabner, Application of Cell-free DNA Analysis to Cancer Treatment, N. Engl. J. Med. 379 (2018) 1754–1765.

[4] N. Fujita, T. Nakayama, N. Yamamoto, S.J. Kim, K. Shimazu, A. Shimosima, N. Maruyama, K. Morimoto, Y. Tamaki, S. Noguchi, Methylated DNA and total DNA in serum detected by one-step methylation-specific PCR is predictive of poor prognosis for breast cancer patients, Oncology 83 (2012) 273–282.

[5] C. Oshiro, N. Kagara, Y. Naoi, M. Shimoda, A. Shimosima, N. Maruyama, K. Shimazu, S.J. Kim, S. Noguchi, PIK3CA mutations in serum DNA are predictive of recurrence in primary breast cancer patients, Breast Cancer Res. Treat. 150 (2015) 299–307.

[6] H. Takahashi, N. Kagara, T. Tani, Y. Naoi, M. Shimoda, A. Shimosima, K. Shimazu, S.J. Kim, S. Noguchi, Correlation of methylated circulating tumor DNA with response to neoadjuvant chemotherapy in breast cancer patients, Clin. Breast Cancer 17 (2017) 61–69 e603.

[7] J.C.M. Wan, C. Mansie, J. Garcia-Corbacho, F. Mouliere, J.D. Brenton, C. Caldas, S. Pacey, R. Baird, N. Rosenfeld, Liquid biopsies come of age: towards implementation of clinical trials using circulating tumor DNA, Transl. Oncol. 32 (2014) 579–586.

[8] X. Zhang, W. Zhao, W. Wei, Z. You, X. Ou, M. Sun, Y. Yin, X. Tang, Z. Zhao, C. Hu, F. Liu, J. Deng, L. Mao, D. Zhou, Y. Ren, X. Li, S. Zhang, C. Liu, J. Geng, G. Yao, B. Song, Y. Liu, D. Li, Y. Ji, Y. Chen, Y. Zhao, S. Yu, D. Pang, Parallel analyses of somatic mutations in plasma circulating tumor DNA (ctDNA) and matched tumor tissues in early-stage breast cancer, Clin. Cancer Res. 25 (2019) 6546–6553.

[9] N. Cancer Genome Atlas, Comprehensive molecular portraits of human breast tumours, Nature 490 (2012) 61–70.

[10] E. Gliud, C. Winter, A. George, Y. Chen, J. Howlin, M.H. Tang, M. Dalgren, R. Schulz, D. Grubau, D. van Westem, M. Forno, C. Ingvar, C. Rose, P. O’Neill, M. Rutherford, L. Ryden, A. Borg, S.K. Gruber-Sash, J. Hjermstom, L.H. Hall, Serial monitoring of circulating tumor DNA in patients with primary breast cancer for detection of occult metastatic disease, EMBIO Mol. Med. 7 (2015) 1034–1047.

[11] G. Schiavon, S. Hrebiien, I. Garcia-Murillas, R.J. Cutts, A. Pearson, N. Tarazona, K. Fenwick, I. Kozarewa, E. Lopez-Knowles, R. Ribas, A. Neriurkur, P. Osin, S. Chandra, L.A. Martin, M. Dowsett, L.E. Smith, N.C. Turner, Analysis of ESR1 mutation in circulating tumor DNA demonstrates evolution during therapy for metastatic breast cancer, Sci. Transl. Med. 7 (2015) 313ra182.