SPANDIDOS PUBLICATIONS

ONCOLOGY LETTERS 15: 7389-7396, 2018

Assessment of basal-like breast cancer by circulating tumor DNA analysis

WEI WEI[1], XIANYU ZHANG[1], SHANSHAN SUN[1], BINGSHU XIA[1], XIAOSHUAN LIANG[1], YAN CUI[1], SONG GAO[1] and DA PANG[1-3]

[1]Department of Breast Surgery, Harbin Medical University Cancer Hospital, Harbin, Heilongjiang 150081; [2]Sino-Russian Medical Research Center; [3]Northern (China) Translational Medicine Research and Cooperation Center, Heilongjiang Academy of Medical Sciences, Harbin, Heilongjiang 150086, P.R. China

Received August 15, 2016; Accepted September 7, 2017

DOI: 10.3892/ol.2018.8224

Abstract. Standardized methods for the detection and assessment of circulating tumor DNA (ctDNA) in breast cancer are not sufficient. In the present study, the method and the potential application of ctDNA in the diagnosis of breast cancer were explored. DNA was extracted from the tumor tissues, plasma and peripheral blood cells of 11 patients with early-stage invasive breast cancer. Primers were designed against the exons of phosphatidylinositol-4,5-biphosphate 3-kinase catalytic subunit α, p53, epidermal growth factor receptor, Akt and phosphatase and tensin homolog. The amplicon-based method for whole-exon sequencing was performed. The associations between the ctDNA mutant frequency with the tumor DNA mutant frequency, and the ctDNA concentration with clinical data were analyzed. A linear association was identified between the concentration of ctDNA and the tumor volume for the 3 patients with basal-like breast cancer, and not in other subtypes. The mutation frequency differed the least between ctDNA and tissue DNA in basal-like breast cancer. ctDNA retained the constituent ratio of gene mutations identified in the corresponding tumor tissue. The ctDNA detection rate depended to a certain extent on the mutation frequency in tumor tissue; for example, a mutant locus with a mutation frequency of >30% in tissues presented a detection rate of >40% in plasma samples, whereas a locus with a mutation frequency of <10% in tissue was associated with a detection rate of ≤1% in the plasma. Therefore, ctDNA may reflect the mutations observed in cancer. Compared with other subtypes, ctDNA may be a more sensitive biomarker for the assessment of mutation and cancer burden in basal-like breast cancer relative to other subtypes.

Correspondence to: Professor Da Pang, Department of Breast Surgery, Harbin Medical University Cancer Hospital, 150 Haping Road, Harbin, Heilongjiang 150081, P.R. China
E-mail: pangdasirhyd@126.com

Abbreviations: ctDNA, circulating tumor DNA; cfDNA, circulating cell-free DNA; SNPs, single-nucleotide polymorphisms; SNVs, single nucleotide variants

Key words: breast cancer, circulating nucleotide, circulating tumor DNA, liquid biopsy

Introduction

At present, the incidence rate for breast cancer is 1/26 in women aged <65 years in China (1). This situation highlights the urgent requirement for improvement in prevention and treatment, making it necessary to search for effective strategies for the early diagnosis and monitoring of breast cancer. Circulating tumor (ct)DNA, which carries tumor-specific sequence alterations, is present in the cell-free fraction of blood, representing a variable, although generally small, fraction of the total circulating cell-free (cf)DNA (2,3). The detection of ctDNA may provide novel strategies for cancer diagnosis and monitoring. The non-invasive analysis of tumor-derived DNA may serve an important role in the identification and analysis of solid tumors, potentially overcoming the limitations of repeated biopsies (4,5). ctDNA exhibits a more significant association with tumor burden than carcinoma antigen 15-3 or circulating tumor cells (6). It has been demonstrated as a potential tool for monitoring metastatic breast cancer with high sensitivity (6). ctDNA may also be used to analyze the mechanisms underlying acquired drug resistance to guide the strategy of tumor management (7). The digital droplet polymerase chain reaction can be applied to monitor ctDNA even in early-stage breast cancer (8). However, the ctDNA detection rate was low in a previous study, due to the limited number of genes and loci analyzed (9). A low detection rate may restrict the feasibility of applying ctDNA analysis.

The aim of the present study was to explore the potential application of ctDNA. Amplicon sequencing was performed on DNA extracted from tumor tissues, plasma and peripheral blood cells in a total of 33 samples from 11 patients with invasive breast cancer. The ratio of gene mutation between ctDNA and tumor DNA was analyzed. In addition, its potential role in monitoring the tumor burden in patients with breast cancer was assessed.

Materials and methods

Sample collection and processing. A total of 11 patients were selected randomly. These patients were all classified as early-stage tumor node metastasis (TNM) stage I/II (10). The tumor tissues were collected through surgical resection of breast tumors and the matched blood samples were collected before surgery. All the specimens were stored at Heilongjiang Province Breast Bio-Sample Bank (Harbin, China). Matched DNA from tumor tissues, plasma and peripheral blood cells were extracted. The patients ranged in age from 36-61, with a mean age of 48.8 years. All patients provided signed informed consent. The study was approved by the Ethical Committee of Harbin Medical University (Harbin, China).

Blood collected in EDTA tubes was processed within 1 h of sample collection and centrifuged at 4˚C, 820 x g for 10 min to separate the plasma from the peripheral blood cells. The plasma was transferred to microcentrifuge tubes and centrifuged at 4˚C, 20,000 x g for 10 min to remove cell debris. Cell-free DNA was extracted from 2-ml aliquots of cell-free plasma using the QIAamp Ultra Sens Virus kit (Qiagen China Co., Ltd., Shanghai, China) according to the manufacturer's protocol. The cell pellet from the initial centrifugation step was used for the isolation of germline genomic DNA using Pure Link Genomic DNA kits (Invitrogen; Thermo Fisher Scientific, Inc., Waltham, MA, USA). Matched tumor DNA was isolated from breast cancer tissues using Pure Link Genomic DNA kits (Invitrogen; Thermo Fisher Scientific, Inc., Waltham, MA, USA).

Sequencing and data analyses. The 5 most commonly mutated genes in breast cancer were identified in The Cancer Genome Atlas (TCGA; http://cancergenome.nih.gov/). The primers used in the amplification of these genes are included in Table I. The amplification products were attached to the Ion AmpliSeq™ Adapters (Life Technologies; Thermo Fisher Scientific, Inc.). Following purification with Agencount Ampure XP Reagent (Beckman Coulter, Indianapolis, IN, USA), the products with Adapters were prepared using the Ion OneTouch™ Template OT2 kit (Life Technologies; Thermo Fisher Scientific, Inc.). The products were then processed using the Ion PGM 200 Sequencing v2 kit (Life Technologies; Thermo Fisher Scientific, Inc.) for sequencing. Sequencing data were analyzed using the Torrent Mapping Alignment program (Torrent Suite™ Software 4.4; Life Technologies; Thermo Fisher Scientific, Inc.). Variants were analyzed in the Torrent variant caller software (Torrent Suite™ Software 4.4; Life Technologies; Thermo Fisher Scientific, Inc.).

Statistical analysis. All analyses were performed using statistical software (SPSS 17.0; SPSS, Inc., Chicago, IL, USA). An unpaired Student's t-test was performed for comparisons of detection rates across molecular subtypes in Table II. A Mann-Whitney test was performed to analyze constituent ratios of gene mutations in Table III. An unpaired Student's t-test was performed to compare the amount of DNA released from tumor tissues into the circulating blood across molecular subtypes in Table IV. Linear-regression analysis was used to evaluate the association between ctDNA concentration and tumor volume in basal-like breast cancer. $P<0.05$ was considered to indicate a statistically significant difference.

Results

Amplicon sequencing. Phosphatidylinositol-4,5-biphosphate 3-kinase catalytic subunit α, TP53, epidermal growth factor receptor, Akt and phosphatase and tensin homolog were identified as the most commonly mutated in breast cancer, according to TCGA. Primers were designed for these genes and the amplicon-based method for whole-exon sequencing was performed on DNA samples from 11 patients with invasive breast cancer from the Heilongjiang Province Breast Bio-Sample Bank, including from tumor tissues, plasma and peripheral blood cells. Ion torrent sequencing was performed; this achieved 98.95% coverage based on a mean sequencing depth of 4,920 reads/sample. The sequenced DNA fragments were 25.2 kb, and the amplicon length was 125-275 bp (coverage >10, quality >100, strand bias <0.8).

Association between the ctDNA mutation rate and clinical data. Subsequent to excluding single-nucleotide polymorphisms (SNPs) detected in the peripheral blood cells, the mean mutation frequencies of the 5 genes for each patient (tested in plasma and tissues) are listed in Table IV. A total of 3/11 patients exhibited the basal-like subtype. Of the patients with the basal-like subtype, the difference in the mutation frequencies between tumor tissues and plasma samples (tissue-plasma) was the lowest in patients 2 and 6, whereas the mutation frequency in plasma was higher than in the tumor tissue of patient 7. Therefore, it was demonstrated that the amount of DNA released from tumor tissues to circulating blood in the basal-like subtype was generally increased compared with that in other subtypes ($P=0.048$). We hypothesized that this was caused by the high proliferative activity and elevated level of necrosis and apoptosis associated with the basal-like subtype (11).

The effect on cell survival and growth of particular mutations lead to an alteration in the mutation compositions between plasma and tumor tissues. To further elucidate this issue, the constituent ratios of gene mutations in plasma and tumor tissues were analyzed (Table III). However, no significant difference in the constituent ratios of gene mutations were observed between plasma and tumor tissues ($P=0.917$, Mann-Whitney test). On account of the similarity in the constituent ratios identified in tumor tissues and plasma in the present study, ctDNA may be useful to track the clonal evolution of tumors.

Comparison of the ctDNA concentration with clinical data. The concentration of ctDNA was previously demonstrated as a biomarker for the assessment of cancer burden. This was verified in the present study. Multiplying the concentration of cfDNA, as determined by a previously determined method (12), by the mutation frequency of ctDNA, yielded the concentration of ctDNA. The mean value for the mutation frequencies of the genes was used as the ctDNA mutation frequency. The clinical data and ctDNA concentrations of the 11 patients are summarized in Table V.

Table I. Primer sequences identified by the Cancer Genome Atlas.

Amplicon_ID	Ion_AmpliSeq_Fwd_Primer*	Ion_AmpliSeq_Rev_Primer*
AMPL7152996431	GGATACGGCCAGGCATTGA	CCCTGTCATCTTCTGTCCCTT
AMPL7152996965	CCCAATTGCAGGTAAAACAGTCA	AGCACTAAGCGAGGTAAGCAAG
AMPL7152996966	AAGCTATGATGTTCCTTAGATTAGGTGTAT	CCAGACCAGCTTTCAAAAAGAAAATTGTTA
AMPL7152996967	TCATAGGGCACCACCACACTAT	CATGGCCATCTACAAGCAGTCA
AMPL7152996968	CATTCTGGGAGCTTCATCTGGA	CGTTCTGGTAAGGACAAGGGTT
AMPL7152996970	TGTGATGAGAGGTGGATGGGTA	CCTCATCTTGGGCCTGTGTTAT
AMPL7152996971	CAGGAAGTCTGAAAGACAAGAGCA	CCAGGGTTGGAAGTGTCTCATG
AMPL7152996972	ATTGCAAGCAAGGGTTCAAAGAC	GCCTTAGGCCCTTCAAAGCATT
AMPL7153002870	AGTGCTTGTTGCTTGCCAG	GGATGTGGACCAACGTGAGG
AMPL7153002873	CCTTCTTGAGCAGCCCTGAAA	CATGTACGAGATGATGTGCGGT
AMPL7153002898	GCGTACTCCATGACAAAGCAGA	GTATCAGGCGACGTGGTATCAA
AMPL7153004003	CGACCCAGTTACCATAGCAATTTAGT	GAAAACTGTTCCAATACATGGAAGGATG
AMPL7153004013	GATCAAGAAATCGATAGCATTTGCAGTATAG	CTCCTAGAATTAAACACACATCACATACAT ACA
AMPL7153004026	CACGACGGGAAGACAAGTTCAT	CAAAACACCTGCAGATCTAATAGAAAACAA
AMPL7153004031	GTTAGCTCATTTTTGTTAATGGTGGCTT	CTCACTCTAACAAGCAGATAACTTTCACTT
AMPL7153004036	TTTTTCTTATTCTGAGGTTATCTTTTTACCACA	CCTCTTGTGCCTTTAAAAATTTGCC
AMPL7153004059	GCCATTTCCATCCTGCAGAAGA	CCAAACTACGGACATTTTCGCATC
AMPL7153007063	TCTTCTCCGAGTGCAGGTAGT	GTTAGGGCTTCTGAGACTTTCCAG
AMPL7153007092	TCAAATGCACCCGAGAAATAAAAACC	CCGTGCTTTAAGAGGTTGGCTT
AMPL7153007099	CACCAGGAAGCCACTCAGATG	GTGGCTATTGTGAAGGAGGGTT
AMPL7153007111	CGCCCTACATGGAAAACCG	CTCACCCAGTGACAACTCAGG
AMPL7153007146	GGCAGGACTCGGCATCAAG	GAACCTCATGCTGGACAAGGA
AMPL7153007158	CCAGTTTTTATCTCCAGCCTCAGTT	AGTGGACCACTGTCATCGAAC
AMPL7153007172	AAATGAGGACCAGGCCAGTTT	GCTTCTTTGCCGGTATCGTGT
AMPL7153007194	CCACCTCGTCCTGTAAAGCAG	CTGCCCATAGACCATGAACGA
AMPL7153009985	CTGTTAGGATCAGATTATAGTGTTACACCA	TGCCATCATTACTTTGATTACAGGATGAT
AMPL7153010057	GGACTATGTCCGGGAACACAAA	ATGGCAAACTCTTGCTATCCCA
AMPL7153010068	ATGTCAGCCGAGGCAGGGAAT	GGGCAAAGAATAAAAGGAAGAGAAAAT CAA
AMPL7153010138	CCACGTACCAGATGGATGTGAAC	GACTCTCCAAGATGGGATACTCCA
AMPL7153011255	GCAAAGACCTGAAGGTATTAACATCATT	AGAGCCAAGCATCATTGAGAAAAGATTTA
AMPL7153011364	ATGTTACTTTTAAAATGAAAAACCTTACA GGAA	ACAGTCCATTGGCAGTTGAGAATAAA
AMPL7153011377	CTTCCAAATCTACAGAGTTCCCTGTT	AGTCCTGTACTTCTGGATCTTTAACCATAT
AMPL7153017580	TGGGATGGTGCTTTGCTGATTA	GTCGAATGTGCTGTTGACACAG
AMPL7153017625	TCGCATCACCAATGCCTTCTTTA	GGCATGTGACAGAACACAGTGA
AMPL7153017648	ATGATTTTTCTTCTCTCCAATGTAGTGGT	GGACCCATTAGAACCAACTCCATAAA
AMPL7153017672	TGCCTGTGGATCCCTAGCTATT	TGATCACCAATCTCTACCAGTTAAAAAGG
AMPL7153017708	ATCGGCCTCTTCATGCGAA	CACATTCAGAGATTCTTTCTGCATCATAAT
AMPL7153017718	ATCCAACATCCAGACACATAGTGATTTT	GAGGACAGTCAGAAATGCAGGAA
AMPL7153017742	ACACCTTCACAATATACCCTCCATGA	CAGGAAGAAAACTCAAGATTATCTGGGTTA
AMPL7153017760	GCAGTTGGGCACTTTTGAAGAT	CCTTTCTCCACTTAGATTTTCTCCAATTTT
AMPL7153017833	GTGTGTCACTCGTAATTAGGTCCA	CGAAAGAAAATACTTGCATGTCAGAGGATA
AMPL7153017844	TCTCTTTAGCAGAACATAAATGCGAAGA	CATAGGAGCTGGAGGCAGAGATA
AMPL7153017850	TTTGAAAGAGAAAAGAAAGAGACATGCATG	TCTCTGAATTTGCAAGAGAGGAAATGTT
AMPL7153017856	GGGAAATGATCCTACCCTCACTCT	CCCACAGGAAGTCTTCTGTCCT
AMPL7153017910	CTGGTAACATCCACCCAGATCA	GCTGCCAGACATGAGAAAAGGT
AMPL7153017922	GACTCCGTCCAGTATTGATCGG	ATCTCAAGGAAACAGGAAAGGACG
AMPL7153017944	TATAATACAGAGTCCCTGAGAGTCTAGAGT	TGCCCTATCTTAGCAACTCTCCT
AMPL7153019495	TTCTTTGTAGATATGATGCAGCCATTGA	CTTGTGATCCAAAAAGTGTCCAAAATCTAT
AMPL7153019513	TACTCCTCTTTCAGAATGTTACCTTATGGT	TGTTCTAACTCAGAGGAATACACAAACAC

Table I. Continued.

Amplicon_ID	Ion_AmpliSeq_Fwd_Primer*	Ion_AmpliSeq_Rev_Primer*
AMPL7153019555	AAGATTGGCCTCCAATCAAACCT	AGCATAAAACTAGTTAGTGCAGTAGGTTTT
AMPL7153019565	AGTTTGGGACTTCTTAAGAAGATTCATATGG	TTTATTTATGTGGACTTTCTGAGAGAAAAC AAT
AMPL7153019574	TCCCTTGAAAAATGAAAGAGAGATGGT	GCAATTACTTGTTCTGGTACACAGTCAT
AMPL7153019590	GCTTTCCTGAAGTTTCTTTTGAAGAGT	GGATAACTTTCAACATACAGGTTGCCTT
AMPL7153019600	TTAGGAATGGATTCCTAAATAAAAATTGAGGT	CAATTCAACCACAGTGGCCTTT
AMPL7153019614	CAGTCTTGCTTCTGTCTCTGAGT	CATGAAATCTGGTCGCCTCATTT
AMPL7153019630	GGGACAACCATACATCTAATTCCTTAAAGT	AAGAAGATTCATCTTGAAGAAGTTGAT GGA
AMPL7153019653	AATATTATGTCTTAGATTGGTTCTTTCCTGTC	GCTGTTCTTTGTCATTTTCCCTTAATTCAT
AMPL7153019708	GTTTGATTACACAGACACTCTAGTATCTGG	ATAAGAGAGAAGGTTTGACTGCCATAA AAA
AMPL7153019717	TTTGGTTGATCTTTGTCTTCGTGATTTG	AAAAATTCAATCAGCGGTATAATCAGGAG
AMPL7153019725	TTCAAAAATGAGTGTTTAAATTGTTTAGCAAAG	TGCTTCAAATACATCCCACATGCA
AMPL7153019740	GTAGTCTGATGTCTCCATTGTTATTAGTGG	AAGGGTTCTCCTCCATGGTAGAT
AMPL7152996430	ACAACCTCCGTCATGTGCT	CTTTCAACTCTGTCTCCTTCCTCTTC
AMPL7152996432	GCTGCCCTGGTAGGTTTTCTG	CGATATTGAACAATGGTTCACTGAAGAC
AMPL7152996963	GAGAATGGAATCCTATGGCTTTCCA	CCGTCATAAAGTCAAACAATTGTAACTTGA
AMPL7152996964	GGGTTATAGGGAGGTCAAATAAGCA	GGCCTCTGATTCCTCACTGATTG
AMPL7152996969	AAGGTGATAAAAGTGAATCTGAGGCAT	AATGGGACAGGTAGGACCTGAT
AMPL7152996973	GGGACTGTAGATGGGTGAAAAGA	AACTGTGAGTGGATCCATTGGAAG
AMPL7152996974	TCAACTTACGACGAGTTTATCAGGAAG	CTTCTTGACTGTTTTACCTGCAATTGG
AMPL7153002838	TTTTTAGGACAAAATGTTTCACTTTTGGG	TTCCTTGTCATTATCTGCACGCT
AMPL7153002851	ACGAACTGGTGTAATGATATGTGCATATTT	CTCAGATCCAGGAAGAGGAAAGGAAAAA
AMPL7153004023	ATCGTTTTTGACAGTTTGACAGTTAAAGG	CGGCTGAGGGAACTCAAAGTAC
AMPL7153004043	GTGTCACATTATAAAGATTCAGGCAATGTT	ATGTATCTCACTCGATAATCTGGATGACT
AMPL7153004047	TTGATTGCTGCATATTTCAGATATTTCTTTCC	TGGCTTAGAAATCTTTTCTAAATGAAAAC ACAA
AMPL7153004054	AAGATGAGTCATATTTGTGGGTTTTCATTTTAA	TTATTTTCATGGTGTTTTATCCCTCTTGAT AAA
AMPL7153007059	CAGCCCGAAGTCTGTGATCTTA	TATGGCGCTGAGATTGTGTCAG
AMPL7153007073	CGCCACAGAGAAGTTGTTGAG	GGGTCTGACGGGTAGAGTGT
AMPL7153007080	CTTCTCATGGTCCTGGTTGTAGAAG	GTCCTGAGCACACGCAATG
AMPL7153007104	GGGTACTAACCTCGTTTGTGCA	CACATTCAGCTTCCTTTGCTTCTC
AMPL7153007116	AGGGACACCTCCATCTCTTCAG	TGGAGCTCCTGATCTGGTACAG
AMPL7153007123	GCCTCTCTGAGTGTGGAGAGAA	TGAAGGTCTTGAGCACACTTGAG
AMPL7153007132	CCTGGAAAGTCTCAGAAGCCCTA	GCCCTGAAGTACTCTTTCCAGA
AMPL7153007163	CTCAGGAGTCTCCACATGGAAG	CTCTGTGAGGCGAGAAACTGAG
AMPL7153007177	GCACCTTCTTCTCGTACACGT	GGCCCTACATCACAGGAGGAA
AMPL7153007184	CTCACGTGCCCAAGAAGACAGGA	CATGAAGATCCTCAAGAAGGAAGTCA
AMPL7153007197	CAGCAGCTTCAGGTACTCAAAC	CAGCCATGTGTTCCCTGTAAG
AMPL7153011276	TCAGAAGTTAAGGCAGTGTTTTAGATGG	GTCCAGAAGTTCCATAGCCTGTTC
AMPL7153011349	ATTTTACAGAGTAACAGACTAGCTAGAGACA	GAAAAAGAAACAGAGAATCTCCATTTTA GCA
AMPL7153017587	CCGAGTATCTCAACACTGTCCA	CTGGATTTTTAGGGCTCATACTATCCTC
AMPL7153017599	AGTCACAGGTTCAGTTGCTTGT	CAGAAGAAATGTTTTTATTCCAAGGGA ACA
AMPL7153017611	CCCAACTCCTTGACCATTACCTC	CCATCATCACTGTTCGGCTTCT
AMPL7153017641	ACCTACCATCATTGGAAAGCAGT	CTGACGACTGCAAGAGAAAACTG
AMPL7153017661	GCATGAACATGACCCTGAATTCG	CAGGAAAATGCTGGCTGACCTA
AMPL7153017684	TGTGATTGTTCATTCATGATCCCACT	GCCTCAGCTGTTTGGCTAAGAG
AMPL7153017693	GCCAAGAAGTGGAATAGCATCTCT	GAGTAGACACAGCTTGAGAGAGAGA

Table I. Continued.

Amplicon_ID	Ion_AmpliSeq_Fwd_Primer*	Ion_AmpliSeq_Rev_Primer*
AMPL7153017702	TTTTTCTCCTTTTAGAAGCTACATAGTGTC	TGCGCTTCCGAACGATGT
AMPL7153017730	CCTCAAAAGAGAAATCACGCATTTATGT	CTGCTGTCTGCATTTATGAACCC
AMPL7153017753	GAGTCTCTGTGTGGAGAGAGTGA	ATCCTCTGGAGGCTGAGAAAATG
AMPL7153017791	AAAGCCATGTTATTCTGCCTTTTTAAACT	GAACAGACACGTGAAGGCATG
AMPL7153017821	GCCTGCCCAACCTACTAATCAG	GGTTGAGGAGCAGGACTGTTTC
AMPL7153017863	GTACGTATTTTGAAACTCAAGATCGCAT	CAGGTACTGGGAGCCAATATTGTC
AMPL7153017872	GTAGCCAGCATGTCTGTGTCA	CCAGAAGGTTGCACTTGTCCA
AMPL7153017881	GGGAGGTCCAGAAAGTGATTGG	CCATATGTCAAATGAATAGTGTCAGGACT
AMPL7153017889	GGTGACCCTTGTCTCTGTGTTC	GGAAATATACAGCTTGCAAGGACTCT
AMPL7153017902	CCACCGTCATCACCTTCCTTTC	GAAGAGGAAGATGTGTTCCTTTGGA
AMPL7153017917	GCCAGCCAAACAATCAGAGAATAAG	CCAATCACCTAAGCAAGTGAAGGA
AMPL7153017929	GCTGGGTTTTCCCACACTAGTG	GGTGGCACCAAAGCTGTATTTG
AMPL7153019490	CTTTGGTCAGGGAACATCTGGA	CAGCACATGAACGTGTAAACAGG
AMPL7153019505	GGAAAGGCAGTAAAGGTCATGCA	CACAGTCACCGATTGACAGACA
AMPL7153019527	CGACAGCATGCCAATCTCTTCA	CCAGAGTGAGCTTTCATTTTCTCAGTTA
AMPL7153019540	TTTCTTTAGATCGGCCATGCAGAA	ACTCTTCCAGCCAAACATAAACAAAAGTATATA
AMPL7153019581	ACAAGCAGAAGTATACTCTGAAAATCA ACC	ACAGATACTCATCCTCAATGTGATTACTTT
AMPL7153019606	GAGTGTCGAATTATGTCCTCTGCAA	AAATGGCTTTCAGTAGTTTTCATGGTTC
AMPL7153019620	TCAGGTACAGATGAAGTTTTTAGTTGAGC	GTAGGAGAATGAGAGAGAGAAGCATAAATT
AMPL7153019634	GCATGAACTATTTAAAGAAGCAAGA AAATACCC	GAATAGGATATTGTATCATACCAATTTCT CGA
AMPL7153019643	TTATGGCAGTCAAACCTTCTCTCTTATG	GGATCCTTTTCCATAGAGAAAGTATCTACC
AMPL7153019670	TGGACACTTTTTGGATCACAAGAAGA	GATTGTTTCTAATAGAGCAGCCAGAACT
AMPL7153019681	ATTTGATGTTGATGGCTAAAGAAAGCC	AAATAATAAGCATCAGCATTTGACTTTACC
AMPL7153019694	GGCATGCCAGTGTGTGAATTTG	ACAGGTAGAAGACTGCACTATAGTAATGAT
AMPL7153019702	GGGAAAAAGGAAAGAATGGGCTT	TGGCCAAAGATTCAAAGCCATTTTT
AMPL7153019732	CCTGCTTTTGGAGTCCTATTGTCG	AAAAGAGTCTCAAACACAAACTAGAGTCA
AMPL7153019748	TTTTTCTTTTAGATCTATGTTCGAACAGGT	CTGGCCAGTGCCTAGCTAATTT

The primers were manufactured specifically for use with the Ion AmpliSeq Kits and contain proprietary modifications.

The scatter plot comparing the ctDNA concentration with tumor volume demonstrated a non-linear association for the majority of the patients (Fig. 1). However, the ctDNA concentration vs. tumor volume demonstrated a highly linear association in the 3 patients with basal-like breast cancer (R^2=0.999, P=0.018). Therefore, ctDNA may provide an effective assessment regarding tumor burden in patients with basal-like breast cancer.

Function of circulating blood cell samples and tumor tissue samples in ctDNA detection. SNPs in the cfDNA were excluded by collectively examining the mutations identified in the blood cell samples. Tumor tissues were used to verify the reliability of ctDNA and exclude the possibility of an unknown source of cfDNA contributing sequence variants that are not SNPs and that do not exist in tumor tissue. SNPs constituted 68% (173/253) of all the single nucleotide variants (SNVs) in cfDNA. Somatic mutations derived from tumor tissues accounted for 26% (65/253). The remaining variants were SNVs of unknown source (15/253, 6%; Table II). As SNPs accounted for most of the SNVs in cfDNA, the detection of germline SNPs should be coupled with ctDNA detection to exclude SNPs from the assessment of cancer burden or prediction of drug sensitivity. There were a limited number of unknown source SNVs, which may have affected the result, even though the number was limited. Therefore, the simultaneous detection of variation between tumor tissue samples and plasma samples was determined to be necessary for accurate data analysis.

Effect of the mutation frequency in tumor tissue on the ctDNA detection rate. The present study identified a similarity in the mutation frequency between tumor tissues and ctDNA in basal-like breast cancer. To determine whether a higher mutation detection rate in ctDNA was observed in patients with basal-like breast cancer, the detection rates across molecular subtypes were compared (Table II). It was demonstrated that basal-like and Her2-positive breast cancer exhibited a markedly increased mutation detection rate compared with the luminal subtype (P=0.033). It was hypothesized that the cause may be the increased frequency of mutations in the two

Table II. SNV component of cfDNA.

	cfDNA SNV number					
Patient	Total SNV	SNP, n (%)	Mutation, n (%)	Unknown source, n (%)	Subtype	P-value
2	29	16 (55)	10 (35)	3 (10)	Basal-like	
6	31	17 (55)	13 (42)	1 (3)	Basal-like	
7	18	16 (89)	0 (0)	2 (11)	Basal-like	
3	27	18 (67)	7 (26)	2 (7)	Her2	0.033
4	24	19 (79)	5 (21)	0 (0)	Her2	
5	22	13 (59)	9 (41)	0 (0)	Her2	
8	23	16 (70)	7 (30)	0 (0)	Luminal A	
1	11	8 (73)	2 (18)	1 (9)	Luminal B $Her2^-$	
9	26	22 (85)	3 (11)	1 (4)	Luminal B $Her2^-$	
10	30	19 (63)	7 (23)	4 (14)	Luminal B $Her2^+$	
11	12	9 (75)	2 (17)	1 (8)	Luminal B $Her2^+$	
Total	253	173 (68)	65 (26)	15 (6)	-	

An unpaired Student's t test was performed (P=0.033). SNV, single nucleotide variants; cfDNA, circulating cell-free DNA; Her2, human epidermal growth factor receptor 2.

Table III. Constituent ratio of gene mutation in plasma and tumor tissue.

	Mutation rate, %		Constituent ratio	
Gene	Plasma	Tissue	Plasma	Tissue
TP53	22.92	28.43	0.27	0.21
Phosphatidylinositol-4,5-biphosphate 3-kinase catalytic subunit α	14.63	21.46	0.17	0.16
Phosphatase and tensin homolog	6.97	17.27	0.08	0.13
Epidermal growth factor receptor	9.79	21.56	0.12	0.16
AKT1	29.63	41.76	0.35	0.32
Total	83.94	130.48	1	1

Table IV. Comparison of the average mutation rate per patient with breast cancer.

	Average mutation rate, %				
Patient	Plasma	Tissue	Tissue-plasma	Subtype	P-value
2	17.69	23.14	5.45	Basal-like	
6	42.11	46.76	4.65	Basal-like	0.048
7	43.07	6.02	-37.05	Basal-like	
3	12.75	20.40	7.65	Her2	
4	8.04	26.53	18.49	Her2	
5	35.00	46.12	11.12	Her2	
8	29.97	42.31	12.34	Luminal A	
9	10.54	19.16	8.62	Luminal B $Her2^-$	
10	21.43	24.94	3.51	Luminal B $Her2^+$	
11	6.38	17.13	10.75	Luminal B $Her2^+$	
1	4.24	22.31	18.07	Luminal B $Her2^-$	

Her2, human epidermal growth factor receptor 2.

molecular subtypes. To further examine this hypothesis, the distribution of mutations derived from tumor tissue in circulating DNA was analyzed (Fig. 2).

A mutant locus with a mutation frequency of >30% in tissues presented a detection rate of >40% in plasma. An additional locus with a <10% mutation frequency in tissue samples exhibited a detection rate of ≤1% in the plasma. A low positive detection rate may be limited by the sensitivity of the detection methods and the number of detected loci. Therefore, the detection rate may be increased by improving the detection sensitivity or increasing the number of detected genes.

Discussion

To the best of our knowledge, the present study is the first to identify a distinction between ctDNA across molecular subtypes. The mutation frequency differed the least between ctDNA and tissue DNA in basal-like breast cancer. Additionally, the association between ctDNA concentration and tumor volume was linear in basal-like breast cancer. Based on these results, ctDNA detection appears to be a

Table V. Clinical data of patients and ctDNA concentration.

Patient	Age, years	Tumor size, cm	Grade	P53 status	TNM	ctDNA (ng/ml)
1	40	2.6x1.3	2	-	IIB	45.79
2	47	1.4x1.2	2	-	IIA	59.79
3	36	1.5x1.1	3	2+	IIA	135.15
4	53	2.0x1.4	2	-	I	26.37
5	42	2.1x2.2	2	2+	IIA	116.20
6	55	2.0x1.6	2	3+	I	513.74
7	43	2.6x1.3	3	-	IIA	615.90
8	60	1.1x1.0	2	-	I	182.81
9	45	2.1x1.7	2	-	IIB	60.49
10	55	2.5x2.0	3	2+	IIB	116.57
11	61	2.0x1.5	1	-	I	19.14

ctDNA, circulating tumor DNA; TNM, tumor node metastasis stage.

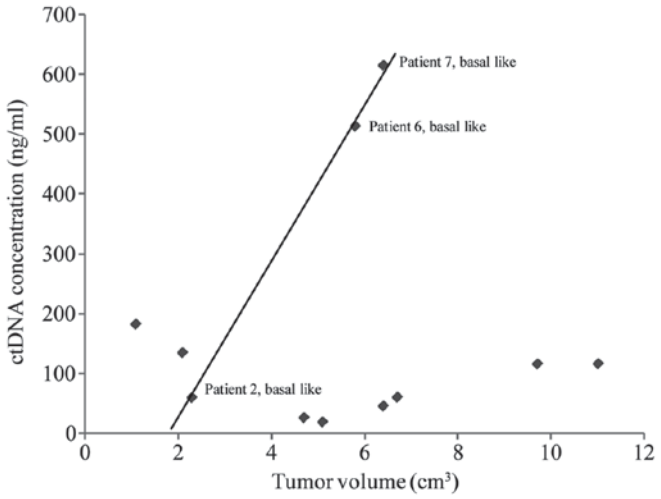

Figure 1. Scatter plot of the correlation between tumor volume and ctDNA concentration. The ctDNA concentration vs. tumor volume exhibited a non-linear association for all tumor types, whereas a linear association in the 3 patients with basal-like breast cancer was observed. ctDNA, circulating tumor DNA.

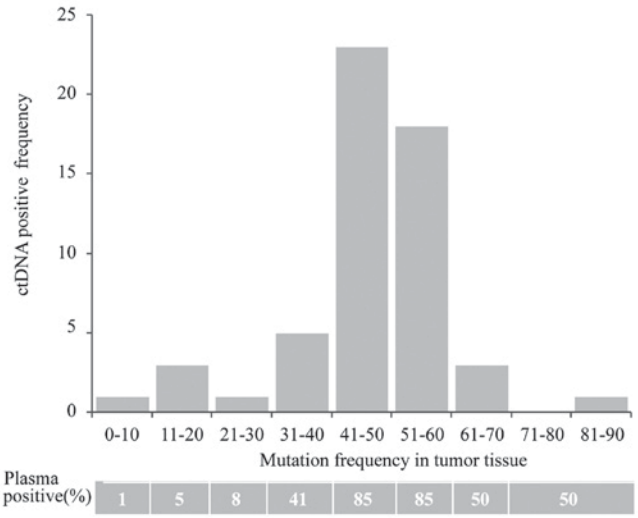

Figure 2. Frequency and positive rate of single nucleotide variants in tumor tissue and plasma. ctDNA, circulating tumor DNA.

promising technique for the assessment of cancer burden in basal-like breast cancer. It has been suggested that ctDNA is a highly sensitive biomarker of metastatic breast cancer (6). However, there is little evidence concerning the significance of tumor burden assessment in primary breast cancer, particularly early-stage breast cancer; the present study demonstrated that ctDNA may allow the assessment of tumor burden in the basal-like subtype of early-stage breast cancer, not in other subtypes.

Understanding the factors that influence the ctDNA detection rate is beneficial to ensure the accuracy of the results. The present study demonstrated that the ctDNA detection rate depended on the mutation frequency in the corresponding tumor tissues. In a previous study, ctDNA was detectable in 50% of samples from patients with localized breast cancer, as assessed by digital PCR of a relatively small number of loci (13). ctDNA was detected in all patients in the present study, due to the greater number of detected loci vs. the previous study. However, the concentration of circulating DNA varied among patients. Therefore, an increased number of mutant loci are required to accurately assess the concentration of ctDNA and improve the reliability of ctDNA measurement in future studies.

In the present study, the mutations identified in tumor tissues exhibited a similar constituent ratio to the mutations identified in plasma. ctDNA may theoretically be used to track tumor subclonal evolution in order to correctly reflect tumor progression. Therefore, liquid biopsy analysis of clonal evolution may be applicable to analyze the origin of cancer. However, SNPs may seriously affect the mutation frequency and concentration of ctDNA as they accounted for almost 70% of all the SNVs in cfDNA in the present study. Therefore, tumor burden assessment required the detection and exclusion of SNPs. In predictions of sensitivity, tissue samples and SNPs should be detected concurrently to accurately identify mutant loci.

In conclusion, the present study indicated that ctDNA may reflect the mutations observed in breast cancer. In basal-like breast cancer, ctDNA may be a particularly sensitive biomarker for the assessment of mutation frequency and cancer burden. Although the detection of circulating DNA was previously reported as a promising method, there are a number of unsolved problems concerning the detection methods, analysis approach and application fields. Previously published ctDNA detection methods employing Tam-Seq (14) or CAPP-Seq (15) may be highly sensitive for the detection and monitoring of ctDNA. However, to further expand the potential clinical applications of ctDNA, additional investigation into the association of mutations with the physical characteristics of each type of cancer is required. The present study suggests that the analysis of ctDNA could be applied clinically to detect and monitor diverse types of malignancy.

Acknowledgements

The authors would like to thank Dr Guangwen Zhang from Heilongjiang Province Breast Bio-Sample Bank (Harbin, China) for assistance with the collection of tissue and blood samples.

Funding

The present study was supported by the Project of Heilongjiang Province Applied Technology Research and Development (grant no., GA13C201).

Availability of data and materials

All data generated or analyzed during this study are included in this published article.

Author's contributions

WW, XZ and DP conceived and designed the study, participated in the analysis and interpretation of the data, and coordinated and drafted the manuscript. SS performed the statistical analysis, participated in the interpretation of the data and assisted in drafting the manuscript. BX participated in the design of the study, statistical analysis, interpretation of the data, coordination and helped in drafting the manuscript. XL participated in the design of the study, interpretation of the data, coordination and assisted in drafting the manuscript. YC and SG participated in the interpretation of the data and assisted in drafting the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate

The study was approved by the Ethical Committee of Harbin Medical University. All patients provided written informed consent.

Consent for publication

All patients provided written informed consent for the publication of any associated data and accompanying images.

Competing interests

The authors declare that they have no competing interests.

References

1. Hao J, Zhao P and Chen WQ: Chines cancer registry annual report, 2012.
2. Schwarzenbach H, Hoon DS and Pantel K: Cell-free nucleic acids as biomarkers in cancer patients. Nat Rev Cancer 11: 426-437, 2011.
3. Gormally E, Caboux E, Vineis P and Hainaut P: Circulating free DNA in plasma or serum as biomarker of carcinogenesis: Practical aspects and biological significance. Mutat Res 635: 105-117, 2007.
4. Whiteside TL: The potential of tumor-derived exosomes for noninvasive cancer monitoring. Expert Rev Mol Diagn 15: 1293-1310, 2015.
5. Nakamura T, Sueoka-Aragane N, Iwanaga K, Sato A, Komiya K, Abe T, Ureshino N, Hayashi S, Hosomi T, Hirai M, *et al*: A noninvasive system for monitoring resistance to epidermal growth factor receptor tyrosine kinase inhibitors with plasma DNA. J Thorac Oncol 6: 1639-1648, 2011.
6. Dawson SJ, Tsui DW, Murtaza M, Biggs H, Rueda OM, Chin SF, Dunning MJ, Gale D, Forshew T, Mahler-Araujo B, *et al*: Analysis of circulating tumor DNA to monitor metastatic breast cancer. N Engl J Med 368: 1199-1209, 2013.
7. Murtaza M, Dawson SJ, Tsui DW, Gale D, Forshew T, Piskorz AM, Parkinson C, Chin SF, Kingsbury Z, Wong AS, *et al*: Non-invasive analysis of acquired resistance to cancer therapy by sequencing of plasma DNA. Nature 497: 108-112, 2013.
8. Beaver JA, Jelovac D, Balukrishna S, Cochran R, Croessmann S, Zabransky DJ, Wong HY, Toro PV, Cidado J, Blair BG, *et al*: Detection of cancer DNA in plasma of patients with early-stage breast cancer. Clin Cancer Res 20: 2643-2650, 2014.
9. Perrone F, Lampis A, Bertan C, Verderio P, Ciniselli CM, Pizzamiglio S, Frattini M, Nucifora M, Molinari F, Gallino G, *et al*: Circulating free DNA in a screening program for early colorectal cancer detection. Tumori 100: 115-121, 2014.
10. NCCN Clinical Practice Guidelines in Oncology. Breast Cancer. Available from: https://www.nccn.org/professionals/physician_gls/pdf/breast.pdf.
11. Rakha EA, Reis-Filho JS and Ellis IO: Impact of basal-like breast carcinoma determination for a more specific therapy. Pathobiology 75: 95-103, 2008.
12. Diehl F, Schmidt K, Choti MA, Romans K, Goodman S, Li M, Thornton K, Agrawal N, Sokoll L, Szabo SA, *et al*: Circulating mutant DNA to assess tumor dynamics. Nat Med 14: 985-990, 2008.
13. Bettegowda C, Sausen M, Leary RJ, Kinde I, Wang Y, Agrawal N, Bartlett BR, Wang H, Luber B, Alani RM, *et al*: Detection of circulating tumor DNA in early- and late-stage human malignancies. Sci Transl Med 6: 224ra24, 2014.
14. Forshew T, Murtaza M, Parkinson C, Gale D, Tsui DW, Kaper F, Dawson SJ, Piskorz AM, Jimenez-Linan M, Bentley D, *et al*: Noninvasive identification and monitoring of cancer mutations by targeted deep sequencing of plasma DNA. Sci Transl Med 4: 136ra68, 2012.
15. Newman AM, Bratman SV, To J, Wynne JF, Eclov NC, Modlin LA, Liu CL, Neal JW, Wakelee HA, Merritt RE, *et al*: An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage. Nat Med 20: 548-554, 2014.