

UCSF

UC San Francisco Previously Published Works

Title

Genomic Profiling of Isolated Circulating Tumor Cells from Metastatic Breast Cancer Patients

Permalink

<https://escholarship.org/uc/item/87w164r5>

Journal

Cancer Research, 73(1)

ISSN

0008-5472

Authors

Magbanua, Mark Jesus M

Sosa, Eduardo V

Roy, Ritu

et al.

Publication Date

2013

DOI

10.1158/0008-5472.can-11-3017

Peer reviewed



Published in final edited form as:

*Cancer Res.* 2013 January 1; 73(1): 30–40. doi:10.1158/0008-5472.CAN-11-3017.

## Genomic profiling of isolated circulating tumor cells from metastatic breast cancer patients

Mark Jesus M. Magbanua<sup>1,2</sup>, Eduardo V. Sosa<sup>1,2</sup>, Ritu Roy<sup>2,3</sup>, Lauren E. Eisenbud<sup>1,2</sup>, Janet H. Scott<sup>1,2</sup>, Adam Olshen<sup>2,4</sup>, Dan Pinkel<sup>2</sup>, Hope Rugo<sup>1,2</sup>, and John W. Park<sup>1,2</sup>

<sup>1</sup>Division of Hematology/Oncology, University of California San Francisco, San Francisco, CA 94143

<sup>2</sup>Helen Diller Family Comprehensive Cancer Center, University of California San Francisco, San Francisco, CA 94143

<sup>3</sup>Helen Diller Family Comprehensive Cancer Center Biostatistics Core, University of California San Francisco, San Francisco, CA 94143

<sup>4</sup>Department of Epidemiology/Biostatistics, University of California San Francisco, San Francisco, CA 94143

### Abstract

Molecular characterization of circulating tumor cells (CTCs) from blood is technically challenging because cells are rare and difficult to isolate. We developed a novel approach to isolate CTCs from blood via immunomagnetic enrichment followed by fluorescence activated cell sorting (IE/FACS). Isolated CTCs were subjected to genome-wide copy number analysis via array comparative genomic hybridization (aCGH). In clinical studies, CTCs were isolated from 181 patients with metastatic breast cancer, 102 of which were successfully profiled, including matched archival primary tumor from five patients. CTCs revealed a wide range of copy number alterations including those previously reported in breast cancer. Comparison with two published aCGH datasets of primary breast tumors revealed similar frequencies of recurrent genomic copy number aberrations. In addition, serial testing of CTCs confirmed reproducibility and indicated genomic change over time. Comparison of CTCs with matched archival primary tumors confirmed shared lineage with notable divergence. We demonstrate that it is feasible to isolate CTCs away from hematopoietic cells with high purity via IE/FACS and profile them via aCGH analysis. Our approach may be utilized to explore genomic events involved in cancer progression and to monitor therapeutic efficacy of targeted therapies in clinical trials in a relatively non-invasive manner.

### Keywords

circulating tumor cells; copy number analysis; array comparative genomic hybridization; genomic instability; metastatic breast cancer

---

Request for reprints: John W. Park, Division of Hematology/Oncology, University of California San Francisco 2340 Sutter St. Box 1387, San Francisco, California 94115 (phone: 415.502.3844; fax: 415.353.9592; jpark@cc.ucsf.edu).

**Conflicts of interest:** JWP: Research Grant (Veridex and BD Biosciences); Honoraria from Speakers Bureau (Veridex); Consultant (Veridex)

## Introduction

Recent studies have demonstrated the clinical significance of detecting circulating tumor cells (CTCs) in blood in breast (1–2) and other cancers (3–4). However little is known about the biology and molecular nature of these cells. The paucity of information can be largely attributed to the technical hurdles in isolating these rare cells (5). Efforts to detect CTCs have involved the use of antibody-based capture methods, thereby greatly enriching the tumor cell component relative to hematopoietic cells (6–9). Current enrichment methods, however, still retain a considerable amount of leukocytes, resulting in a heterogeneous admixture that is difficult to analyze (10–11). Attempts to isolate individual tumor cells can be technically challenging (12–13).

Here we describe the feasibility of a novel strategy comprising immunomagnetic enrichment and fluorescence activated cell sorting (referred to as IE/FACS), followed by genomic profiling of the isolated CTCs. In the initial step, magnetic beads coated with EpCAM mAb are used to enrich for tumor cells and then subjected to FACS analysis using differentially labeled mAbs to distinguish tumor cells (EpCAM+) from leukocytes (CD45+) during sorting. Using IE/FACS, we isolated samples of highly purified CTCs from breast cancer patients. Genomic DNA was then subjected to whole genome amplification (WGA) followed by array comparative genomic hybridization (aCGH) to determine genome-wide copy number aberrations in CTCs. The assay was evaluated using breast tumor cell model systems and then applied to blood samples from 181 metastatic breast cancer (MBC) patients, 102 of whom were successfully profiled.

## Materials and Methods

### Cell lines

Control (MCF7, BT474) cancer cell lines were obtained from American Type Culture Collection (ATCC, Manassas, VA) and verified using polyphasic (genotypic and phenotypic) testing to confirm identity. Cell lines were passaged in our laboratory for less than six months and grown in the supplier recommended complete growth media supplemented (DME-H21, RPMI-1640) with 10% fetal calf serum, 50µg/ml streptomycin sulfate and 50 U/ml penicillin G. Cells were incubated at 37°C in a humidified incubator with 5% CO<sub>2</sub>. BT474 and MCF7 breast cancer cell lines were chosen for proof-of-principle experiments because they have numerous and well characterized copy number changes and focal amplifications (14–15). Trypsinized cultured BT474 (10,000), MCF7 (1,000) cells were spiked into 10mL of blood from a healthy donor and tumor cells were isolated by IE/FACS (see below).

### Patients and samples

Clinical samples were obtained from MBC patients who were recruited to participate in institutional (UCSF) or multicenter cooperative group trials (Cancer and Leukemia Group B (CALGB) 40502 and 40503, Translational Breast Cancer Research Consortium (TBCRC) 009; Supplementary Table 1). All patients gave informed consent under a protocol approved by the Institutional Review Board in respective centers participating in the specific clinical

trials. Approximately 10mLs of blood were drawn into CellSearch™ cell save preservative tubes (Veridex) for CTC enumeration and additional 10–20mLs were drawn into tubes containing EDTA for CTC isolation. Blood samples from other institutions were shipped overnight and processed immediately upon receipt. CTCs in 7.5 mLs of blood were enumerated using CellSearch™.

### Primary tumor samples

Microdissection of archival formalin-fixed paraffin-embedded (FFPE) primary tumor were done as previously described (16). Areas microdissected contained at least 80% tumor. Microdissected tissue was collected in CellLytic™ cell lysis buffer and proteinase K (4:1 ratio; Sigma) and incubated at 60°C for 1 hr and 95°C for 4 min. Genomic DNA present in the whole cell lysate (1–2 µL) was subjected to whole genome amplification followed by aCGH analysis of the amplified DNA (see below).

### Cell isolation via IE/FACS

To increase the likelihood of isolating CTCs for genomic analysis, we enumerated CTCs in 7.5mL using CellSearch™ to identify patients with  $\geq 1$  CTC/mL. CTCs were then isolated from the remaining blood samples by IE/FACS. First, whole blood was diluted by adding one-half volume of Cell Buffer (BD Bioscience), a proprietary sodium azide-free blocking reagent that helps minimize non-specific binding of antibodies; alternatively, Dilution Buffer (Veridex) was also used in place of the Cell Buffer with comparable results. Next, EpCAM (MJ37) mAb-coated magnetic beads (50 µg/10mL of indirect conjugate or 60µg/10mL for the direct conjugate) and EpCAM (EBA-1) mAb conjugated to phycoerythrin (PE) (400µL at 5µg/mL per 10mL of blood) were added to the diluted blood. The sample was incubated for 15 minutes, mixing half-way and again at the end. Magnetic field was applied by placing the sample in a magnetic separator (Immunicon) for 45 minutes. Following application of the magnetic field, red blood cells were aspirated without lysis along with unbound cells. The bound cells were washed with 2mL of Cell Buffer, then moved to a 12 × 75 mm tube, and subjected again to the magnetic field for 5 minutes. The cells were then aspirated and resuspended in 150µL of Cell Buffer. Cells were stained prior to FACS analysis by adding 20µL of a solution containing 2µg of a proprietary nucleic acid dye (BD Biosciences), and 0.1µg of a leukocyte-specific CD45 (2D1) mAb conjugated to peridinin-chlorophyll-protein-Cy5.5. After 15-minute incubation in the dark, 450uL 1X phosphate buffered saline was added. Tumor cells were sorted using FACS Aria (BD Biosciences) into PCR tubes containing 10µL of TE (10mM Tris-HCl 1mM EDTA pH 8.0). We defined CTCs as nucleated (nuclear stain-positive), EpCAM-positive, and CD45-negative. All the steps in IE/FACS were performed at room temperature. Sources and details regarding IE/FACS antibodies and reagents are described in Supplementary Table 2.

### Whole genome amplification (WGA)

Genomic DNA from lysed CTCs was subjected to WGA using the GenomePlex Single Cell WGA4 Kit (Sigma-Aldrich) following manufacturer's protocol. To prevent loss of genomic material, we performed cell lysis and subsequent WGA of genomic DNA in the same PCR tube the cells were sorted. Briefly, we added 1 µL lysis and proteinase K digestion solution to cells in 10µL of TE. Then, DNA fragmentation and library preparation was performed.

Next, 7.5µL of 10x Amplification Master Mix, 48.5µL of nuclease-free water and 5µL WGA DNA polymerase were added to the resulting library. To amplify the library, the samples were subjected to an initial denaturation of 95°C for 3 min followed by 25 cycle of a denaturation step at 94°C for 30s and an annealing/extension step at 65°C for 5 min. QIAQuick PCR Purification Kit (Qiagen) was used to clean up the amplified DNA and concentration was determined by a NanoDrop™ spectrophotometer and then stored at -20°C.

### Array comparative genomic hybridization (aCGH) analysis

ACGH analysis was performed as previously described (17). Quality control of aCGH data was performed by calculating the median absolute deviation (MAD) and choosing the cut-off of 0.20 (see Supplementary Fig. 1). All data are MIAME compliant, and the raw data has been deposited into GEO under super-series accession # GSE27511. See Supplementary Information for details of aCGH analysis.

### Copy number analysis

The microarray data was subjected to circular binary segmentation (CBS) (18) as implemented in the DNACopy package from Bioconductor (19) to translate intensity measurements into regions of equal copy number and to make gain/loss/amplification calls (see Supplementary Information for details). CBS represents an heuristic approach based on empirical distribution of the data (18). There is no established method to calculate confidence thresholds for CBS. Enrichment tests were done at the arm level to identify significantly gained and lost chromosome arms. The Fridlyand (20) and Chin (21) cohort data were processed in a similar way. Eight of the samples in the Chin cohort were excluded from the analysis as they were already part of the Fridlyand cohort. Frequency plots of copy number alterations were based on clones that were in common among cohorts.

Concordance between array CGH profiles was measured using weighted Pearson correlation. A weighted Pearson measure was used because, in CGH comparisons, unweighted Pearson correlation is driven by the number of ratios that differ from “normal.” Since the observations in this case are the clones, one can use the spatial relationship of the clones to estimate the weights. Based on this idea, the weight of a clone was estimated by considering its deviation and that of its adjacent clone from “normal”. The formula used for calculating weighted correlation coefficient (corr) is given below:

$$r_w = \frac{\text{cov}(x, y; w)}{\sqrt{\text{cov}(x, x; w)\text{cov}(y, y; w)}}$$

where x and y are the log2 ratio values of the two samples under consideration; w is the

weight vector;  $w_i = \frac{1}{2}(\text{median}(|x_j - x_{ref}|) + \text{median}(|y_j - y_{ref}|))$ ;  $i = 1, 2, \dots, n-1$ ;  $j = i, i+1$ ;  $n$  = total number of clones; and  $x_{ref}$  and  $y_{ref}$  are the medians of the n clones for samples x and y, respectively. Only autosomal clones were considered.

The levels of correlation between sample pairs of primary tumors and CTCs in our cohort and the Fridlyand cohort were set by evaluating the distribution of weighted correlation coefficients of sample pairs in the Chin cohort. Correlation coefficients falling in the percentile intervals 0 to < 50, 50 to < 75, 75 to < 99 and 99 to 100 of the Chin cohort were said to be un-, lowly-, moderately- and highly- correlated respectively. For this study, the corresponding quintiles were 0 to < 0.18, 0.18 to < 0.29, 0.29 to < 0.65 and 0.65 to 1. For cell line data, a separate set of 27 samples consisting of 21 BT474, 4 PC3 and 2 LNCaP cell line replicates were used to evaluate the correlation coefficients of cell line sample pairs in the same way. Here, the same percentile intervals as above were designated as being un-, lowly-, moderately- and highly- correlated were from 0 to < 0.80, 0.80 to < 0.88, 0.88 to < 0.98 and 0.98 to 1 respectively.

The overall concordance of gain between cohorts was measured by estimating the concordance correlation coefficient of the proportion of gains. The concordance correlation coefficient (22) for the pair of cohorts represented as  $x$  and  $y$  was calculated as

$$r_c = \frac{2s_{xy}}{s_x^2 + s_y^2 + (\bar{x} - \bar{y})^2}, \text{ where the mean for } x \text{ is } \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \text{ the variance for } x \text{ is}$$

$s_x^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$ , the covariance between  $x$  and  $y$  is  $s_{xy} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$ ,  $x_i$  is the proportion of samples gained in clone  $i$  for  $x$  and  $n$  is the number of clones. Only autosomal clones were considered. The overall concordance of lost clones between cohorts was measured similarly. Statistical analyses were performed using the R language (23).

We used the Nexus 6.0 (Biodiscovery Inc.) copy number analysis program to perform an exploratory analysis comparing genomic aberrations prevalent in CTCs versus primary tumors. Nexus 6.0 performs a Fishers Exact test statistical comparison to determine the likelihood of having an observed event frequency, e.g. specific losses or gains, in one group vs. another group based on random chance. Gains and losses with 30% difference in frequency between the two datasets with  $p$  value and  $q$ -bound value (corrected for multiple testing) < 0.05 were considered statistically significant.

## Results

### Methods development and validation

Cells recovered by IE/FACS appeared to be tumor cells without any detectable leukocyte contamination as confirmed by multiple assessments. For example, direct microscopy of isolated cells showed tumor cell morphology exclusively (see Supplementary Information and Supplementary Fig. 2). FISH assay of isolated cells showed known copy number changes in spiked BT474 cells, and characteristic breast cancer abnormalities in patient samples (see Supplementary Information and Supplementary Fig. 3). Conversely, leukocytes separately collected as a negative control exhibited a normal FISH result. Importantly, aCGH of isolated cells from spike experiments indicated no detectable admixture from contaminating leukocytes (see below).

Breast cancer cells from *in vitro* culture were spiked into healthy donor blood and then recovered via IE/FACS. Isolated tumor cells were then subjected to WGA and analyzed by aCGH. Genomic profiling of IE/FACS-isolated BT474 tumor cells (range, 1–50 cells) demonstrated extensive copy number changes (Fig. 1A), as did aCGH of unamplified BT474 DNA performed as a positive control (Fig. 1B). Both aCGH profiles were fully consistent with published results for this cell line (14–15). Comparison of the aCGH profiles of IE/FACS-isolated cells vs. positive control BT474 DNA showed good correlation (Table 1). These results suggest minimal dilution from contaminating normal leukocyte DNA, which would have otherwise dampened the dynamic range of copy number alterations (14). Genomic analysis was possible with small numbers of cells, including single cells (Supplementary Fig. 4). Similarly, MCF7 cells spiked into healthy blood were successfully isolated via IE/FACS and profiled via WGA and aCGH (Supplementary Fig. 5); the resulting aCGH profile showed the expected genomic aberrations reported for MCF7 (15).

### Genomic profiling of CTCs from breast cancer patients

We next tested the feasibility of this approach in clinical samples. CTCs were isolated from a series of metastatic breast cancer (MBC) patients via IE/FACS and subjected to WGA/aCGH. Enrichment tests were done at the chromosome arm level to identify significantly gained and lost chromosome arms (Supplementary Information). A complete list of gained and lost arms for each sample is found in Supplementary Table 3.

Genomic profiling of 112 CTCs from patient #302 with ER positive, HER2 negative MBC showed gains in several chromosome arms including focal amplifications at 8p11–12, 8q24, 11q13, 20q13, and losses including 8p and 13q (Fig. 2A). Gain at 11q13 is commonly observed in hormone receptor positive breast cancer (24), and is also frequently co-amplified with 8p12 (25) as observed here.

Profiling of 20 CTCs from patient #22 with triple negative MBC revealed gains in several chromosome arms including 3q, 10p, 12p, 21q, and focal amplifications at 8q24 and 13q31. Losses were observed including those on 1p, 5p and 8p (Fig. 2B). High level gain at 13q31 has been previously reported, particularly in association with triple negative breast cancer (24).

In patient #108 with ER positive HER2 negative MBC, only 7 CTCs were isolated. aCGH profiling found notable aberrations, including gains in 8q and 17q, with focal amplification in 15q26; and losses in 3p, 8p and 11q (Fig. 2C).

### Genomic profiling of CTCs from an expanded cohort

We prospectively evaluated this approach for CTC isolation and profiling in a larger series of MBC patients, including patients enrolled in multicenter cooperative group clinical trials. We screened 181 consecutive patients to reach our goal of ~100 informative CTC profiles, as follows: 181 patients were found to have 1 CTC/mL (range 1.0–202.7 CTC/mL, Supplementary Table 1); we isolated ~20 CTCs from each patient via IE/FACS (range: 4–250 cells) and performed WGA followed by aCGH analysis. Of the 181 samples, 16 (9%) samples resulted in low WGA yield, 52 (29%) samples did not pass WGA QC testing and 11

(6%) samples resulted in aCGH profiles with high background noise (Supplementary Information and Supplementary Fig. 6). The remaining 102 (56%) samples produced informative aCGH data and were used for subsequent analysis.

Examination of the genomic profiles of CTCs from 102 patients revealed a wide range of copy number aberrations. CTC samples were enriched with low level alterations across the genome that included gains in 1q and 8q and losses in 1p, 2q, 4q, 8p, 11q, 13q, 15q, 16q, 18q (Fig. 3A). Comparison between CTCs in this study and breast primary tumors (N=62) in a previously published aCGH study by Fridlyand et al (20) revealed a similar frequency distribution of copy number gains and losses (concordance correlation coefficient,  $r_c = 0.68$  for both gains and losses) (Fig. 3B). To confirm this agreement, we estimated the concordance of our CTC data to another published aCGH dataset of 137 primary breast tumors by Chin (21) et al (Fig. 3C). This comparison again revealed high concordance ( $r_c = 0.72$  and  $0.73$  for gain and loss respectively), while concordance between the two primary tumor datasets was higher ( $r_c = 0.90$  and  $0.80$  for gain and loss respectively). Frequent focal amplifications reported in both Fridlyand and Chin datasets e.g., on 8p (including *FGFR1*), 8q (*MYC*), 11q13 (including *CCND1*), 17q (*ERBB2*) and regions on 20q (including *ZNF217*) were also observed in CTCs [11%, 11%, 12%, 5%, 5%, respectively].

To determine genomic aberrations prevalent in CTCs, we performed an exploratory comparative analysis between our CTC CGH dataset and primary tumor CGH from the Fridlyand et al (20) dataset. This comparison suggested that gains on 5q13 (including *CCNB1*), 7q22 (including *MUC12* and *MUC17*), 9p13 and 9q31 and losses on 10q22 and 8p23 were significantly more frequent in CTCs compared to primary tumors (Supplementary Table 4). However, since CTC and primary tumor samples were not processed in the same manner, this analysis must be regarded as exploratory.

This study was not designed to develop a clinical diagnostic assay nor to find clinical correlates, and thus we did not explore associations with clinicopathological variables or clinical outcomes. However, we note that aCGH of CTCs from 30 (29%) patients showed copy number gain at the 17q (HER2) locus, and 5 (5%) of these showed high level gains (focal amplification). Information on clinical HER2 status was available for only 10 of the 102 patients. Two primary tumors were known to be HER2 positive, and both patients had CTCs also showing focal amplification of HER2. Eight primary tumors were HER2-negative, and the respective CTC samples showed six with no HER2 copy number gain and two with low-level gain in HER2 copy number.

### Assay performance characteristics

We tested whether the number of cells as input for WGA, the level of CTCs in the blood (CTC/mL) and the interval (days) between blood draw and IE/FACS isolation were associated with successful WGA and aCGH analysis. Quality control (QC) measures at each step are discussed in detail in the Supplementary Information. As expected, a larger number of cells subjected to WGA was significantly associated with the likelihood of successfully passing WGA QC (p-value <0.001). Higher CTC levels (CTC/mL) were also associated with successful WGA QC (p-value= 0.001). 74% of the samples were processed within 1 day post-phlebotomy, while 18% and 8% were processed within 2 or 3 days, respectively.



The interval from blood draw to CTC isolation was not associated with a significant difference in WGA QC outcome. No significant associations were observed for any of the parameters tested and likelihood of successfully passing the aCGH QC.

### Serial genomic profiling of CTCs

CTCs were isolated at multiple time points in three patients (#303, see next section for #462 and #4013) to evaluate the feasibility of serial sampling and profiling of CTCs.

Patient #303 is a 50-year old female with ER/PR positive, HER2 negative MBC. On Day 1, the patient had 2.4 CTC/mL and 74 CTCs were isolated by IE/FACS. Genomic profiling of CTCs obtained on Day 1 revealed aberrations including loss of 8p and gain of 8q (Fig. 4A). On Day 128, her CTC count increased to 9.9 CTC/mL, and 50 CTCs were isolated. On Day 169, her CTC count further increased to 94.9 CTC/mL, and three replicates of 50 CTCs each were isolated. Profiling of CTCs isolated on Days 128 and 169, respectively, revealed numerous additional copy number aberrations beyond the baseline Day 1 profile, including focal amplifications in 6q21, 8q24, and 19q13 and losses of several chromosome arms including 3p, 4p, 8p, 11q and 13q (Fig. 4B and 4C). In addition, correlation of the three sorting replicates at Day 169 was high (mean  $r_w$  0.87, sd 0.03) indicating reproducibility of the assay (Supplementary Fig. 7). Correlation analysis between the baseline (Day 1) and 2<sup>nd</sup> or 3<sup>rd</sup> (Day 128 or 169) time points revealed low concordance (mean  $r_w$  0.46, sd 0.06), while correlation of the two later time points (Day 128 vs. Day 169) revealed high concordance (mean  $r_w$  0.91, sd 0.03); these results suggest major clonal shift occurring between the baseline and 2<sup>nd</sup> time point. Of note, the patient initially responded to her cancer treatment with resolution of ascites and pleural effusion, but subsequently developed clinical disease progression which was documented after the final sample on Day 169.

### Genomic profiling of CTCs versus matched primary tumor

In five patients for whom CTC isolation and profiling were performed, the corresponding archival primary tumor tissue was retrieved and analyzed by aCGH as well.

Patient #4009 is a 41-year old female with triple negative breast cancer with rapid metastatic relapse after adjuvant chemotherapy. 500 CTCs were isolated from the patient's blood by IE/FACS. After cell lysis, the sample was divided into halves and the replicates subjected to whole genome amplification in parallel. aCGH analysis revealed copy number alterations including gains in 7q, 8q and losses in 8p and 13q (Fig. 5A). Amplification replicates showed high correlation ( $r_w=0.93$ ), demonstrating the reproducibility of the WGA/aCGH protocol. The patient's archived primary tumor from 2.5 years prior was later obtained and analyzed by aCGH. Comparison of the profiles of the CTCs and primary tumor, respectively, showed conserved genomic aberrations, including gains in 8q, 10p, and 14q and losses in 4q, 8p, 13q and 15q. Gain in 10p has been previously reported to be prevalent in triple negative breast cancer (24). Overall, the CTC and primary tumor profiles showed high correlation (mean  $r_w$  0.80, sd 0.01), consistent with a clonal relationship.

Patient #4013 is a 50-year old female with ER/PR positive, HER2 positive MBC. Twenty CTCs were isolated at two different time points (Days 1 and 42). Both CTC samples

revealed focal amplifications on 8q24, 12q15, 17q12 (*HER2*) and 20q13 (Fig. 5B). Correlation analysis between the two profiles revealed high concordance ( $r_w=0.92$ ). The archived primary tumor from six years prior was then obtained. aCGH analysis showed multiple aberrations in common with the CTC samples, including the same focal amplifications. Overall, the primary tumor and CTC profiles were highly correlated (mean  $r_w$  0.89, sd 0.03). However, some aberrations, e.g. losses in 6q, 13q, 18q and 20p, were observed only in CTCs. Also, more low level alterations overall were observed in CTCs compared to the primary tumor, suggesting that the CTC samples were of higher purity (lesser contaminating normal DNA) or that CTCs had acquired additional alterations.

Patient #462 is a 63-year old female with ER/PR negative and HER2 equivocal MBC. Her primary tumor was reportedly HER2 positive by FISH testing (HER2/CEP17 ratio = 3.1) at her local hospital, but subsequent biopsy of one of the bone metastasis at UCSF was HER2 negative by FISH. Retesting of her primary tumor at UCSF showed equivocal HER2 IHC but negative amplification by FISH (ratio = 1.5). The patient initially had bone only metastatic disease. Subsequently, she was noted to have progressive disease involving new lung metastasis and increased bone metastasis. Her CTC level at that time (Day 1) was 59.7 CTC/mL, and CTCs were isolated. The initial anti-EpCAM-enriched cell population was subjected to two independent FACS isolations, which yielded two sorting replicates of 20 CTCs each. The patient continued to have disease progression; on Day 87 her CTC level further increased to 137.6 CTC/mL and a new sample of 20 CTCs was isolated. Genomic DNA from each sample was independently amplified and analyzed for copy number aberrations (Fig. 5C) CTC samples exhibited low level alterations across the genome, including loss in 1p, 8p, 11q, 16q, 18q. Correlation analysis revealed high concordance (mean  $r_w$  0.68, sd 0.03) among the three CTC profiles. The archived primary tumor from six years prior was then obtained and profiled by aCGH. The primary tumor showed aberrations in common with the CTCs, including losses in 8p, 11q and 16q and gain in 7p. Interestingly, the primary tumor, which had yielded conflicting clinical HER2 results, showed a low level gain at the HER2 locus, and the CTCs showed no evidence of HER2 amplification. Overall, the primary tumor profile showed moderate correlation with the CTC profiles (mean  $r_w$  0.51, sd 0.08).

Patient #4015 is a 54-year old female with triple negative breast cancer. Twenty CTCs were isolated from the patient's blood by IE/FACS. aCGH analysis showed focal amplification on 8q24, gains in 10p and 19q, and losses in 3p, 5q and 6q (Supplementary Fig. 8A). The original primary tumor and a lymph node metastasis from 2.5 years prior were then obtained and subjected to aCGH analysis. aCGH profiles of the primary tumor and nodal metastasis, respectively, were highly correlated with each other ( $r_w=0.87$ ), but only moderately correlated with the CTC profile (mean  $r_w$  0.46, sd 0.01). Overall, CTCs exhibited additional genomic aberrations beyond those observed in the primary tumor and nodal metastasis, such as copy number gain in 20q and loss in 3p.

Patient #4026 is a 59-year old female with ER/PR negative and HER2 positive breast cancer. Seven (7) CTCs were isolated from her blood, and aCGH analysis revealed high level copy number gain in 17q12 (*HER2*) as well as focal amplifications in 6p12 and 6q22. aCGH analysis of the primary tumor from 5 years prior also demonstrated the expected

HER2 amplification on 17q12 (Supplementary Fig. 8B, arrow); however, the additional copy number aberrations (e.g. focal amplifications on chromosome 6) observed in CTCs were clearly absent in the archival primary tumor. A moderate correlation ( $r_w=0.46$ ) between the CTC and primary tumor was observed.

Taken together, these results indicate a clear clonal relationship between primary tumors and subsequent CTCs, and the appearance of new, as well as conserved genomic alterations. Exploratory analysis in these 5 matched cases suggested copy number changes associated with CTCs but not primary tumors, but these were not statistically significant after correction for multiple comparisons due to the small sample size.

## Discussion

We have developed a novel approach for molecular profiling of CTCs via sequential immunomagnetic enrichment and FACS sorting to isolate CTCs, followed by whole genome DNA amplification and array CGH analysis. For methods validation, BT474 and MCF7 cells spiked into normal blood were isolated and correctly profiled. This approach was then applied to clinical samples, including a series of 102 CTC samples which were successfully profiled by aCGH. Immunomagnetic enrichment/FACS provided robust CTC isolation and enabled detailed molecular analysis. Of note, our assay utilizes two different mAbs against the cell surface marker EpCAM, thus obviating the need for cell permeabilization, e.g. immunocytochemical detection of cytokeratin, which may affect the suitability of cells for downstream analyses. Additionally, high concordance among technical replicates confirmed the reproducibility of the assay.

Our approach involving immunomagnetic enrichment, FACS sorting and aCGH analysis is novel, and this report confirms the feasibility of this approach in an extended series of metastatic breast cancer patients. To our knowledge, this is the first demonstration of aCGH analysis of CTCs isolated from breast cancer patients, as well as the largest dataset on genome-wide copy number analysis of any CTCs. While other studies have reported copy number analysis of CTCs in other cancer types, our approach is also unique because it involves full isolation of CTCs; this contrasts with analyses of mixed cell populations enriched for CTCs via cell adherence (26) or immunomagnetic separation alone (27).

Many CTC detection strategies rely on EpCAM expression; the detected circulating epithelial cells (CECs) are then presumed to be malignant. This is a reasonable assumption since detection of CECs in non-cancer patient samples is extremely rare (1). However, it remains a hypothetical possibility that CECs might include circulating non-malignant epithelial cells. In our study, the EpCAM-positive cells isolated via IE/FACS exhibited a wide range of copy number aberrations, unequivocally indicating their malignant origin. Moreover, these CTCs showed breast cancer-associated genomic alterations frequently seen in primary breast tumors; and selected patients with both CTCs and archival primary tumor showed clonal lineage. We therefore conclude that in metastatic breast cancer patients, CECs are generally CTCs. We also note that aCGH profiles of CTCs did not appear to be significantly dampened by normal DNA, either from non malignant epithelial cells or

leukocytes. Indeed, we view isolation of CTCs without significant leukocyte contamination as a particular advantage of the IE/FACS approach.

Genomic profiling of CTCs revealed numerous copy number alterations, including many previously reported in primary breast tumors. Frequent copy number aberrations identified in our series of 102 CTC samples included gains in 1q and 8q and losses in 1p, 2q, 4q, 8p, 11q, 13q, 15q, 16q, 18q. Focal amplifications included 8p11–12 (*FGFR1*), 8q24 (*MYC*), 11q13 (*CNND1*), 17q12 (*HER2*), and 20q13 (*ZNF217*) (20–21, 28).

We describe the first longitudinal analysis of CTC copy number alterations using multiple samples over time in the same patients. Results of these analyses further confirmed the reproducibility of this approach, indicating conservation of genomic alterations over time in individual patients. In addition, there was evidence for emergence of new alterations in the course of disease progression. For example, CTCs from patient #303 exhibited new genomic alterations over time, such as high level copy gains on 5q, 6q, 12p and 19q that were not present in CTCs isolated at an earlier time point (Fig. 4).

This is also to our knowledge the first report demonstrating a clonal relationship between breast CTCs and matched archival primary tumors via genome-wide copy number analysis. CTCs were demonstrated to have both conserved and divergent genomic alterations as compared to the corresponding primary tumor. These exploratory comparisons will need to be further studied in a larger cohort of matched primary tumors and CTCs. For example, it will be of interest whether candidate copy number changes suggested by the initial analysis of our CTC cohort as compared to an independent primary tumor cohort will be confirmed in matched cases of primary tumor vs. CTCs. Also, this IE/FACS approach can in principle be used in conjunction with other profiling strategies, such as expression analysis or DNA sequencing; such strategies may further elucidate molecular features particularly associated with CTCs rather than primary tumors.

In summary, we demonstrate an approach to isolate CTCs away from hematopoietic cells with high purity via IE/FACS and to profile them via WGA and aCGH analysis. Genomic profiling confirmed the malignant nature of these cells. In addition, we showed that copy number aberrations in CTCs reflect those known to occur in breast primary tumors. This approach may be used to explore genomic events associated with breast cancer progression, and may provide more detailed CTC-based biomarkers in clinical trials.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

**Funding:** This work was supported by The National Cancer Institute (NCI) CA31946 awarded to the Cancer and Leukemia Group B, NCI U54 CA90788, NCI Early Detection Research Network (EDRN) U01 CA111234, NCI SP0RE P50 CA58207, The Avon Foundation for Women, The Breast Cancer Research Foundation, and The Susan G. Komen for the Cure for the Translational Breast Cancer Research Consortium (TBCRC).

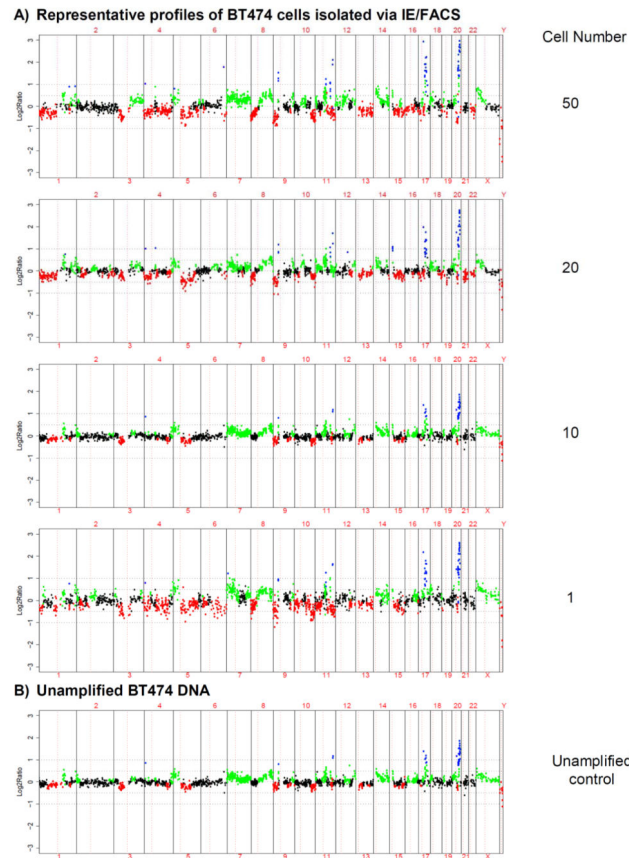
We thank Michele Melisko and Mark Moasser for providing samples; Alfred Au for review of pathology slides; Stephanie Cohen, Artem Rysantzev, Yesenia Ramos, Erin Bowlby, members of the Helen Diller Family

Comprehensive Cancer Center Laboratory for Cell Analysis Core (Sarah Elmes and Jane Gordon) and the Array Core (Greg Hamilton) for technical assistance; the Cancer and Leukemia Group B (CALGB, Monica M. Bertagnoli, Chair) for funding and providing most of the patient samples (CALGB 40502 and 40503) and the CALGB Statistical Center (Daniel J. Sargent and William Barry) and Breast Correlative Science Committee for critical review of the manuscript.

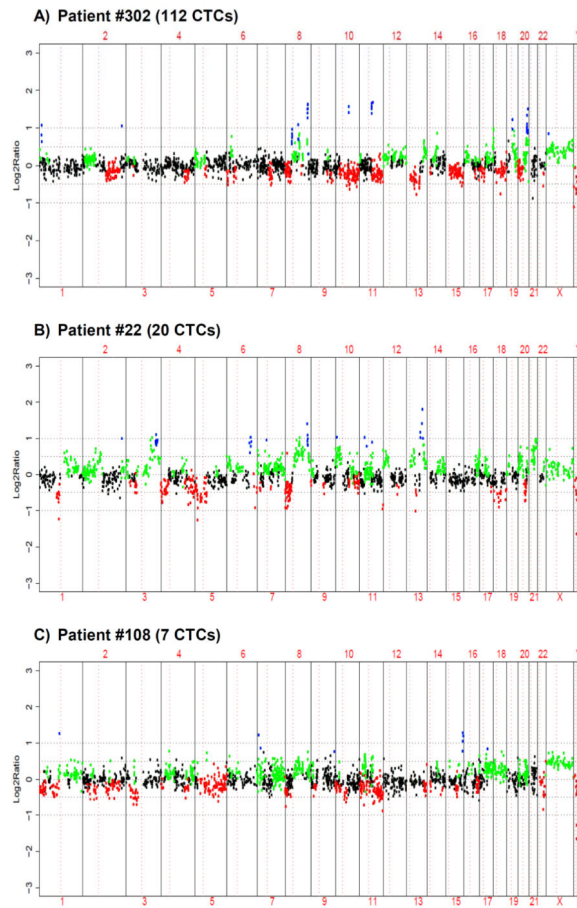
## References

1. Cristofanilli M, Budd GT, Ellis MJ, Stopeck A, Matera J, Miller MC, et al. Circulating tumor cells, disease progression, and survival in metastatic breast cancer. *N Engl J Med.* 2004; 351:781–791. [PubMed: 15317891]
2. Liu MC, Shields PG, Warren RD, Cohen P, Wilkinson M, Ottaviano YL, et al. Circulating tumor cells: a useful predictor of treatment efficacy in metastatic breast cancer. *J Clin Oncol.* 2009; 27:5153–5159. [PubMed: 19752342]
3. Cohen SJ, Punt CJ, Iannotti N, Saidman BH, Sabbath KD, Gabrail NY, et al. Relationship of circulating tumor cells to tumor response, progression-free survival, and overall survival in patients with metastatic colorectal cancer. *J Clin Oncol.* 2008; 26:3213–3221. [PubMed: 18591556]
4. Danila DC, Heller G, Gignac GA, Gonzalez-Espinoza R, Anand A, Tanaka E, et al. Circulating tumor cell number and prognosis in progressive castration-resistant prostate cancer. *Clin Cancer Res.* 2007; 13:7053–7058. [PubMed: 18056182]
5. Pantel K, Brakenhoff RH, Brandt B. Detection, clinical relevance and specific biological properties of disseminating tumour cells. *Nat Rev Cancer.* 2008; 8:329–340. [PubMed: 18404148]
6. Nagrath S, Sequist LV, Maheswaran S, Bell DW, Irimia D, Ulkus L, et al. Isolation of rare circulating tumour cells in cancer patients by microchip technology. *Nature.* 2007; 450:1235–1239. [PubMed: 18097410]
7. Racila E, Euhus D, Weiss AJ, Rao C, McConnell J, Terstappen LW, et al. Detection and characterization of carcinoma cells in the blood. *Proc Natl Acad Sci U S A.* 1998; 95:4589–4594. [PubMed: 9539782]
8. Talasz AH, Powell AA, Huber DE, Berbee JG, Roh KH, Yu W, et al. Isolating highly enriched populations of circulating epithelial cells and other rare cells from blood using a magnetic sweeper device. *Proc Natl Acad Sci U S A.* 2009; 106:3970–3975. [PubMed: 19234122]
9. Stott SL, Lee RJ, Nagrath S, Yu M, Miyamoto DT, Ulkus L, et al. Isolation and characterization of circulating tumor cells from patients with localized and metastatic prostate cancer. *Sci Transl Med.* 2010; 2:25ra23.
10. Smirnov DA, Zweitzig DR, Foulk BW, Miller MC, Doyle GV, Pienta KJ, et al. Global gene expression profiling of circulating tumor cells. *Cancer Res.* 2005; 65:4993–4997. [PubMed: 15958538]
11. Sieuwerts AM, Kraan J, Bolt-de Vries J, van der Spoel P, Mostert B, Martens JW, et al. Molecular characterization of circulating tumor cells in large quantities of contaminating leukocytes by a multiplex real-time PCR. *Breast Cancer Res Treat.* 2009; 118:455–468. [PubMed: 19115104]
12. Geigl JB, Speicher MR. Single-cell isolation from cell suspensions and whole genome amplification from single cells to provide templates for CGH analysis. *Nat Protoc.* 2007; 2:3173–3184. [PubMed: 18079717]
13. Klein CA, Schmidt-Kittler O, Schardt JA, Pantel K, Speicher MR, Riethmuller G. Comparative genomic hybridization, loss of heterozygosity, and DNA sequence analysis of single cells. *Proc Natl Acad Sci U S A.* 1999; 96:4494–4499. [PubMed: 10200290]
14. Pinkel D, Segraves R, Sudar D, Clark S, Poole I, Kowbel D, et al. High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. *Nat Genet.* 1998; 20:207–211. [PubMed: 9771718]
15. Neve RM, Chin K, Fridlyand J, Yeh J, Baehner FL, Fevr T, et al. A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer Cell.* 2006; 10:515–527. [PubMed: 17157791]
16. Waldman FM, DeVries S, Chew KL, Moore DH 2nd, Kerlikowske K, Ljung BM. Chromosomal alterations in ductal carcinomas in situ and their in situ recurrences. *J Natl Cancer Inst.* 2000; 92:313–320. [PubMed: 10675380]

17. Snijders AM, Nowak N, Segreaves R, Blackwood S, Brown N, Conroy J, et al. Assembly of microarrays for genome-wide measurement of DNA copy number. *Nat Genet.* 2001; 29:263–264. [PubMed: 11687795]
18. Olshen AB, Venkatraman ES, Lucito R, Wigler M. Circular binary segmentation for the analysis of array-based DNA copy number data. *Biostatistics.* 2004; 5:557–572. [PubMed: 15475419]
19. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, et al. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* 2004; 5:R80. [PubMed: 15461798]
20. Fridlyand J, Snijders AM, Ylstra B, Li H, Olshen A, Segreaves R, et al. Breast tumor copy number aberration phenotypes and genomic instability. *BMC Cancer.* 2006; 6:96. [PubMed: 16620391]
21. Chin K, DeVries S, Fridlyand J, Spellman PT, Roydasgupta R, Kuo WL, et al. Genomic and transcriptional aberrations linked to breast cancer pathophysiologies. *Cancer Cell.* 2006; 10:529–541. [PubMed: 17157792]
22. Lin LI. A concordance correlation coefficient to evaluate reproducibility. *Biometrics.* 1989; 45:255–268. [PubMed: 2720055]
23. R Development Core Team R. A language and environment for statistical computing. Vienna, Austria: 2010.
24. Han W, Jung EM, Cho J, Lee JW, Hwang KT, Yang SJ, et al. DNA copy number alterations and expression of relevant genes in triple-negative breast cancer. *Genes Chromosomes Cancer.* 2008; 47:490–499. [PubMed: 18314908]
25. Kwek SS, Roy R, Zhou H, Climent J, Martinez-Climent JA, Fridlyand J, et al. Co-amplified genes at 8p12 and 11q13 in breast tumors cooperate with two major pathways in oncogenesis. *Oncogene.* 2009; 28:1892–1903. [PubMed: 19330026]
26. Paris PL, Kobayashi Y, Zhao Q, Zeng W, Sridharan S, Fan T, et al. Functional phenotyping and genotyping of circulating tumor cells from patients with castration resistant prostate cancer. *Cancer Lett.* 2009; 277:164–173. [PubMed: 19162393]
27. Ulmer A, Schmidt-Kittler O, Fischer J, Ellwanger U, Rassner G, Riethmuller G, et al. Immunomagnetic enrichment, genomic characterization, and prognostic impact of circulating melanoma cells. *Clin Cancer Res.* 2004; 10:531–537. [PubMed: 14760074]
28. Chin SF, Teschendorff AE, Marioni JC, Wang Y, Barbosa-Morais NL, Thorne NP, et al. High-resolution aCGH and expression profiling identifies a novel genomic subtype of ER negative breast cancer. *Genome Biol.* 2007; 8:R215. [PubMed: 17925008]

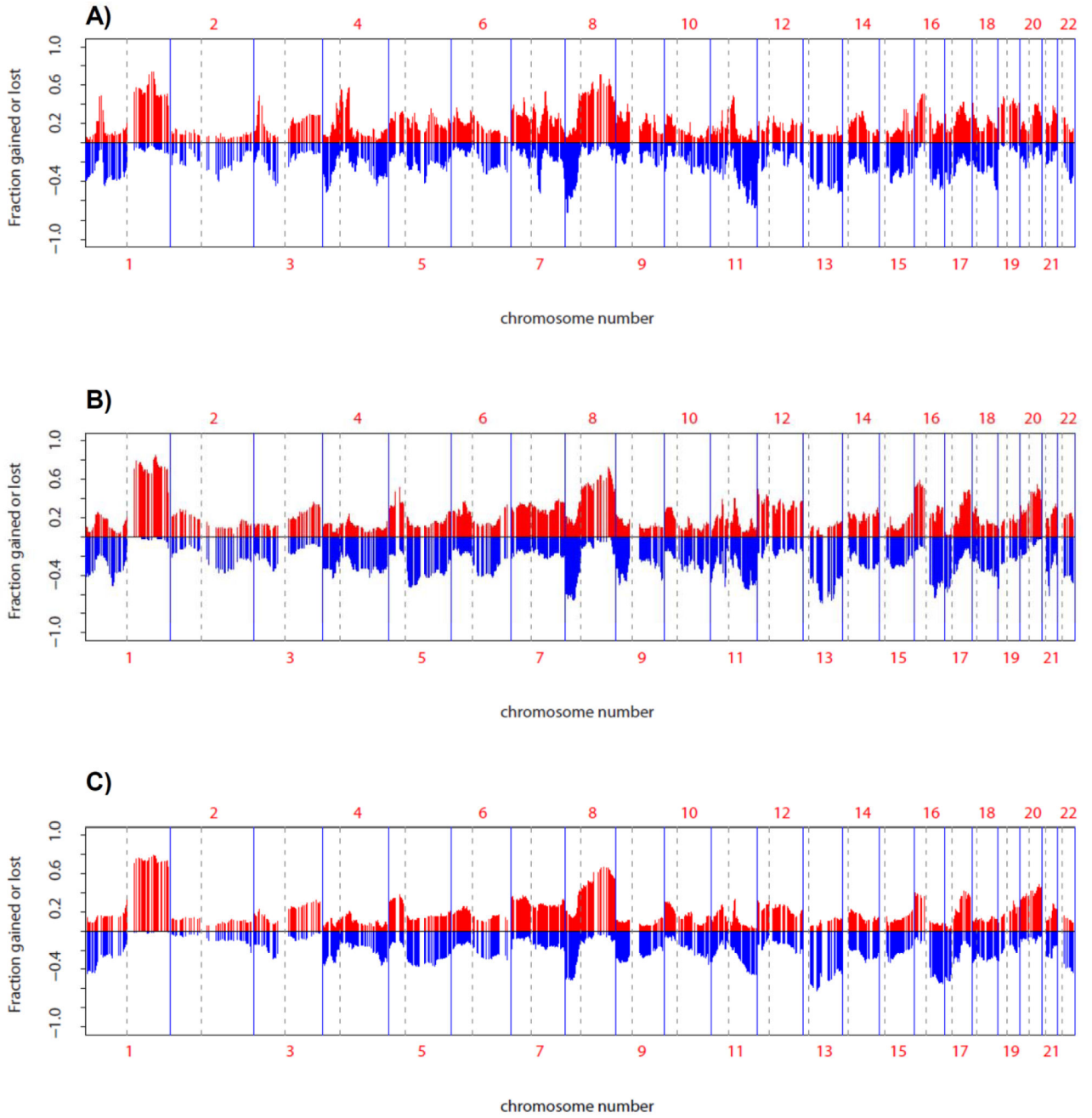
**Fig. 1.**

Copy number analysis of BT474 cells isolated via immunomagnetic enrichment and FACS (IE/FACS). (A) Representative ACGH profiles generated from 1 to 50 BT474 cells and (B) unamplified BT474 DNA positive control. The  $\log_2$  ratio value for each BAC clone is plotted on the y-axis. The x-axis represents the genomic position of each BAC clone on the array, with chromosome numbers indicated. Vertical solid lines indicate chromosome boundaries, and vertical red dashed line represents the centromeric region dividing each chromosome into the p- or short arm (to the left of centromere) and the q- or long arm (to the right of the centromere). Color represents copy number status: red=loss, green=gain, blue=amplification, and black=no change.

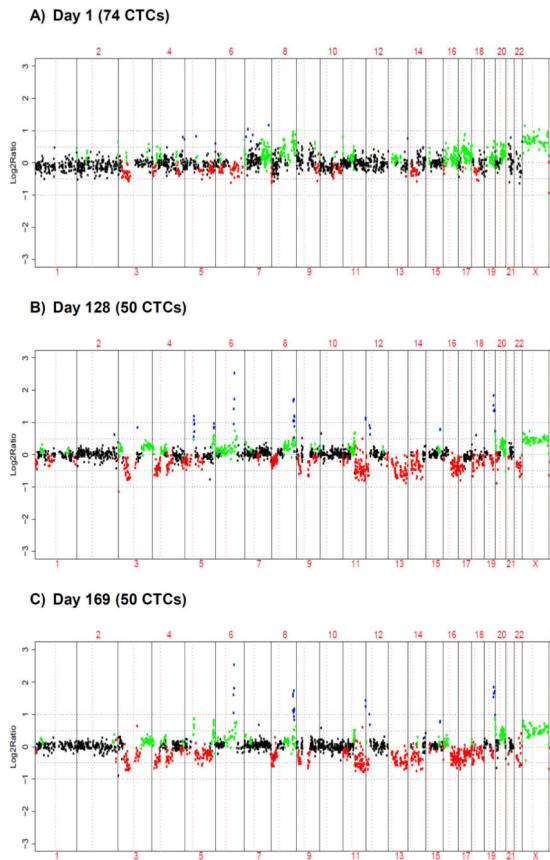


**Fig. 2.** Copy number analysis of CTCs. Genomic profiles from (A) 112 CTCs from patient #302 (B) 20 CTCs from patient #22 and (C) 7 CTCs from patient #108.

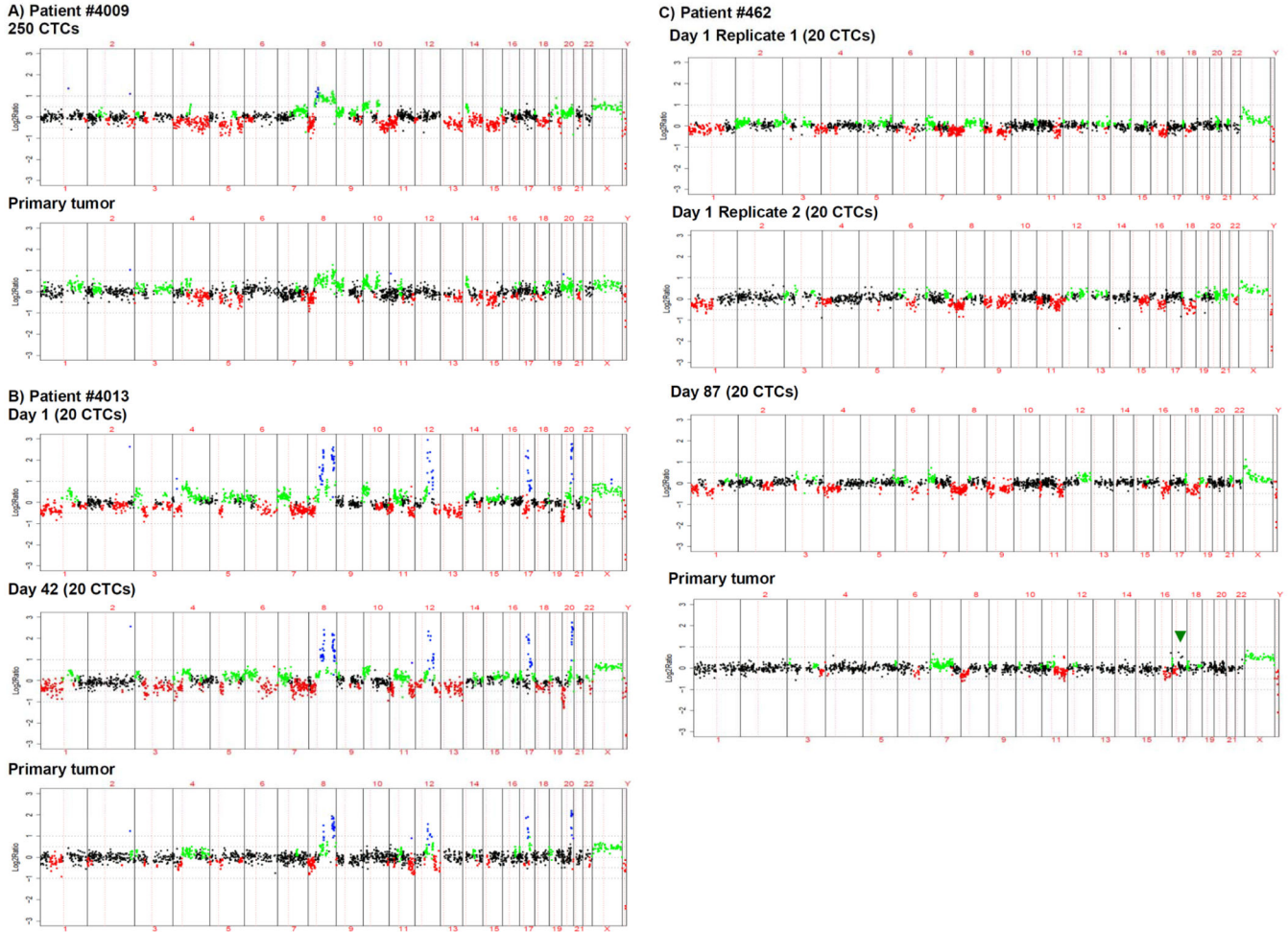




**Fig. 3.** Recurrent copy number aberrations in CTCs. Frequency plot of copy number alterations in (A) CTCs from 102 metastatic breast cancer patients, (B) in 62 primary breast tumors (Fridlyand et al 2006 (20)) and (C) 137 primary breast tumors (Chin et al 2006 (21)). Gains and losses are shown in red and blue, respectively.



**Fig. 4.** Serial copy number analysis of CTCs. Genomic profiles of CTCs from patient #303. (A) 74 CTCs at day 1. (B) 50 CTCs at day 128. (C) 50 CTCs at day 169 (Profiles from triplicate of 50 CTCs sorted independently from same blood draw are shown in Supplementary Figure 5).



**Fig. 5.** Copy number analysis of CTCs and matched primary tumor. (A) Profiles from 250 CTCs and the primary tumor of patient #4009. (B) Profiles from 20 CTCs isolated from two independent blood draws (days 1 and 42) and the primary tumor of patient #4013. (C) Genomic profiles from sorting replicates of 20 CTCs at day 1 isolated from the same blood sample, then amplified and arrayed independently as well as 20 CTCs at day 87 and the primary tumor of patient #462. The primary tumor showed a low level gain in HER2 (arrow).

**Table 1**  
Correlation analysis of genomic profiles of spiked BT474 isolated from blood via immunomagnetic enrichment and FACS (IE/FACS). Isolated 1 to 1000 BT474 cells were subjected to aCGH analysis and compared to a profile generated from unamplified BT474 genomic DNA from cell culture (positive control). See Figure 1 and Supplementary Figure 2 for representative aCGH profiles.

No. of BT474 cells	Correlation Coefficients within group			Correlation Coefficients of each group with unamplified sample		
	Mean correlation	SD of correlation	No. of samples	Mean correlation	SD of correlation	No. of samples (including unamplified)
1	0.93	0.02	3	0.86	0.01	4
10	0.97	NA	2	0.91	0.01	3
20	0.90	NA	2	0.93	0.02	3
50	0.87	0.02	4	0.89	0.04	5
100	0.94	0.02	3	0.94	0.03	4
500	0.98	NA	2	0.94	0.00	3
1000	NA	NA	1	0.93	NA	2