Analysis of tumor template from multiple compartments in a blood sample provides complementary access to peripheral tumor biomarkers

William M. Strauss[1], Chris Carter[1], Jill Simmons[1], Erich Klem[1], Nathan Goodman[2], Behrad Vahidi[1], Juan Romero[1,8], Michael Masterman-Smith[1], Ruth O'Regan[3], Keerthi Gogineni[4], Lee Schwartzberg[5], Laura K. Austin[6], Paul W. Dempsey[1], Massimo Cristofanilli[7]

[1]Cynvenio Biosystems, Westlake Village, CA, 91361, USA

[2]Independent, Seattle, WA, 98109, USA

[3]Department of Hematology and Medical Oncology, Winship Cancer Center, Emory University, Atlanta, GA, 30322, USA

[4]Division of Hematology/Oncology, University of Wisconsin, Madison, WI, 53792, USA

[5]The West Clinic, Germantown, TN, 38138, USA

[6]Department of Medical Oncology, Kimmel Cancer Center, Thomas Jefferson University, Philadelphia, PA, 19107, USA

[7]Lurie Cancer Center, Northwestern University, Chicago, IL, 60611, USA

[8]Current Address: Xencor, Inc, Monrovia, CA, 91016, USA

Correspondence to: *Paul W. Dempsey,* ***e-mail:*** *pdempsey@cynvenio.com*

Keywords: *liquid biopsy, CTC, cfDNA, metastatic breast cancer, next generation sequence*

Received: *December 23, 2015* ***Accepted:*** *March 14, 2016* ***Published:*** *March 30, 2016*

ABSTRACT

Targeted cancer therapeutics are promised to have a major impact on cancer treatment and survival. Successful application of these novel treatments requires a molecular definition of a patient's disease typically achieved through the use of tissue biopsies. Alternatively, allowing longitudinal monitoring, biomarkers derived from blood, isolated either from circulating tumor cell derived DNA (ctcDNA) or circulating cell-free tumor DNA (ccfDNA) may be evaluated. In order to use blood derived templates for mutational profiling in clinical decisions, it is essential to understand the different template qualities and how they compare to biopsy derived template DNA as both blood-based templates are rare and distinct from the gold-standard. Using a next generation re-sequencing strategy, concordance of the mutational spectrum was evaluated in 32 patient-matched ctcDNA and ccfDNA templates with comparison to tissue biopsy derived DNA template. Different CTC antibody capture systems for DNA isolation from patient blood samples were also compared. Significant overlap was observed between ctcDNA, ccfDNA and tissue derived templates. Interestingly, if the results of ctcDNA and ccfDNA template sequencing were combined, productive samples showed similar detection frequency (56% vs 58%), were temporally flexible, and were complementary both to each other and the gold standard. These observations justify the use of a multiple template approach to the liquid biopsy, where germline, ctcDNA, and ccfDNA templates are employed for clinical diagnostic purposes and open a path to comprehensive blood derived biomarker access.

INTRODUCTION

Cancer remains one of the leading causes of morbidity worldwide. Treatment decisions and response monitoring is historically dependent on serial imaging technologies and disease-specific pathologic characterization of tissue biopsies typically obtained at time of primary surgery. This generic and standard approach does not support effective cures in the advanced setting primarily because of the inability of cytotoxic therapy to

deal with tumor heterogeneity. In the last few years, the emergence of next generation sequencing tools has cast a very different light on the nature of tumor clonality with the ability to build an emerging model describing a much more heterogeneous disease than previously understood [1, 2]. This raises problems in determining the best diagnostic approach when dealing with a disease characterized by a dynamic plasticity that could be not captured in its complexity by the simple molecular snapshot offered by a one-time tissue biopsy. The traditional biopsy is increasingly understood as too restrictive to monitor relevant changes in progression and resistance because it represents a "geographically" and "temporally" restricted sample with implications in patients' management [3]. This has driven the development of various "liquid biopsy" technologies that seek to address the need for monitoring tools that use the specific readout of DNA based biomarkers to monitor changes in tumor profile [4].

The earliest version of a liquid biopsy was built on careful enumeration of a small number of cells that could be found in the blood of patients with epithelial cancers [5, 6]. In the most mature form, this resulted in the FDA approved CellSearch™ test. Using EpCAM to recover, and intracellular expression of Cytokeratin to detect a population of cells, it was shown this population of circulating epithelial cells predict worse outcome, faster disease progression, and increased likelihood of metastatic events [7]. This population came to be known as circulating tumor cells (CTC). However there were several problems demonstrated with the approach. Molecular evaluation has shown some but not all of the cells bear molecular hallmarks of cancer [8-10]. Also, this restricted phenotypic definition of a tumor cell was described prognostically and so excluded many classes of informative cells [11-14]. As a result, CTC enumeration using these legacy definitions frequently returns limited or absent numbers of cells. Maybe most importantly, CTC counting has been ineffective at demonstrating clinical utility in the advanced setting for individual patients [15]. Despite those limitations, because of their detection in a peripheral blood sample, circulating tumor cells provide an attractive source of genetic material for longitudinal monitoring in view of the minimal invasiveness of a blood draw and their potential to reflect the molecular profile of the metastatic cell population [8, 16-17].

Circulating cell-free DNA was identified in 1948 in the plasma and derives from both normal and diseased tissue [18]. Clinical studies have shown ccfDNA can act as a suitable template for cancer monitoring and management [19-22]. Unlike circulating tumor cells, studies suggest ccfDNA representation aligns with disease burden and disease type [23]. As such, the reasons ctcDNA and ccfDNA are found in blood are different with CTC representing the mobile cellular aspect of a tumor while ccfDNA is produced chiefly as a product of apoptosis [24, 25].

In summary, two primary sources of DNA template are available from cancer patient blood samples. These templates have been interrogated for specific known mutations using ultra sensitive PCR based techniques [26]. However the mutational spectrum from these two sample types have never been compared because of the signal to noise challenge for enriching rare tumor cells and a clear comparison of template quality has never been performed. Furthermore, there has not been a concordance study between ctcDNA, ccfDNA, and formalin fixed, paraffin embedded tissue (FFPE) from the same patient and blood draw using sequencing. One reason for this has been technical. Previous studies to compare the two templates have reached different conclusions based on technical limitations for enrichment techniques or based on evaluation with molecular tools that are not directly comparable [21]. Recent advances in separation technologies have provided a solution for this technical problem [27]. The LiquidBiopsy® platform (Cynvenio Biosystems), allows for isolation of useful amounts of ccfDNA and ctcDNA from the same patient blood draw. Typically the amount of ccfDNA and ctcDNA recovered per blood draw is sufficient to produce patient matched NGS libraries. The use of NGS means tumor cells and tumor derived DNA fragments can be defined as mutation bearing events and can therefore be directly compared. In this report we describe a generalizable strategy using NGS on 32 matched FFPE, ccfDNA and ctcDNA samples from clinical samples.

RESULTS

Characterization of clinical samples

In this study, patients were recruited based on a confirmed diagnosis of metastatic breast cancer. The samples were collected at baseline, either before start of a new therapy or at completion of staging diagnosis. A tissue biopsy specimen of the metastatic recurrence was mandatory. The patients were predominantly female (97%) and all had stage IV disease. The remaining characteristics are shown in Table 1. Formalin-fixed, paraffin-embedded (FFPE) tumor specimens were obtained for all patients for whom sufficient tissue remained in the pathology block of the metastatic lesion. Of 32 patients, 7 had either insufficient DNA recovered for sequencing, too little or no tumor component to the biopsy or no remaining biopsy tissue. 25 of 32 samples were successfully evaluated.

Target recovery and extraction of ctcDNA and ccfDNA template

EpCAM is the legacy marker for recovery of CTC in breast cancer. To expand the definition of cells that can be functionally recovered, a cocktail of antibodies that recognize surface receptors in addition to EpCAM

Table 1: Characteristics of normal and metastatic cancer samples analyzed for peripheral templates

Characteristic		Number	Range (%)
Normal Controls			
Number			
	total Normals	31	
Age (years)			
	Median	48	30-74
Sex			
	Male	13	(42%)
	Female	15	(48%)
	Unknown	3	(10%)
Metastatic Cancer Samples			
Number			
	Total	32	
Age (years)			
	Median	55	36-82
Sex			
	Male	1	(3%)
	Female	31	(97%)
BC subtype			
	Luminal A	15	(47%)
	Luminal B	6	(19%)
	Her2	4	(13%)
	TNBC	6	(19%)
	Unknown	1	(3%)
Pathologic stage			
	Stage IV	31	(97%)
	Unknown	1	(3%)
No. of lines of Therapy			
	Median	4	0-13
Sampling Interval			
	Synchronous (≤6 months)	17	0-18 weeks (53%)
	Asynchronous (≥6 months)	14	31-240 weeks (44%)
	Not available	3	(3%)

were evaluated. Tumor cells representing the five major subtypes of breast cancer were added to 7.5mL normal healthy blood samples and recovered using the LiquidBiopsy system (Figure 1). The five subtypes; Basal, Claudin low, Her2, Luminal A and Luminal B, were each represented by two different well characterized tumor cell lines [28]. The efficiency of recovery using EpCAM was compared to capture with EpCAM, Her2, and Trop2 (an epithelial/mesenchymal transition profile (EMT)). Recovery (top graph) and purity (bottom graph) of samples engineered with 90c/mL of each cell line were determined. The same cell lines were evaluated by

FACs for expression of the three markers. EpCAM only based recovery gave between <1% and 103% recovery of engineered samples. Consistent with patterns of expression, Claudin low and Basal type breast cancer cell lines demonstrated the least recovery with EpCAM alone. Addition of Her2 and Trop2 specific antibodies incremented recovery of target cells when those receptors were expressed. So HCC-1569, MDA-MD-231 and HCC-1937 cell recovery was significantly improved by the pool of antibodies. There is a consistent inhibition of capture of MCF7 cells when the mixture of antibodies is used possibly due to stearic hinderance. Despite this, when a receptor is available for capture, the 90c/mL samples were recovered with an average efficiency of 77% using the EMT cocktail. Critically, with the exception of HCC-1395 which demonstrated no significant capture with either reagent set, 90 tumor cells/mL could be enriched to an average purity of between 29% and 67% from whole blood (see bottom graph). Thus, the LiquidBiopsy platform could reproducibly enrich target cells on the order of 10^5-10^6 fold enrichment.

Applying this approach to clinical samples, tumor cell populations were enriched using EpCAM targeted capture and compared to EMT targeted capture. This

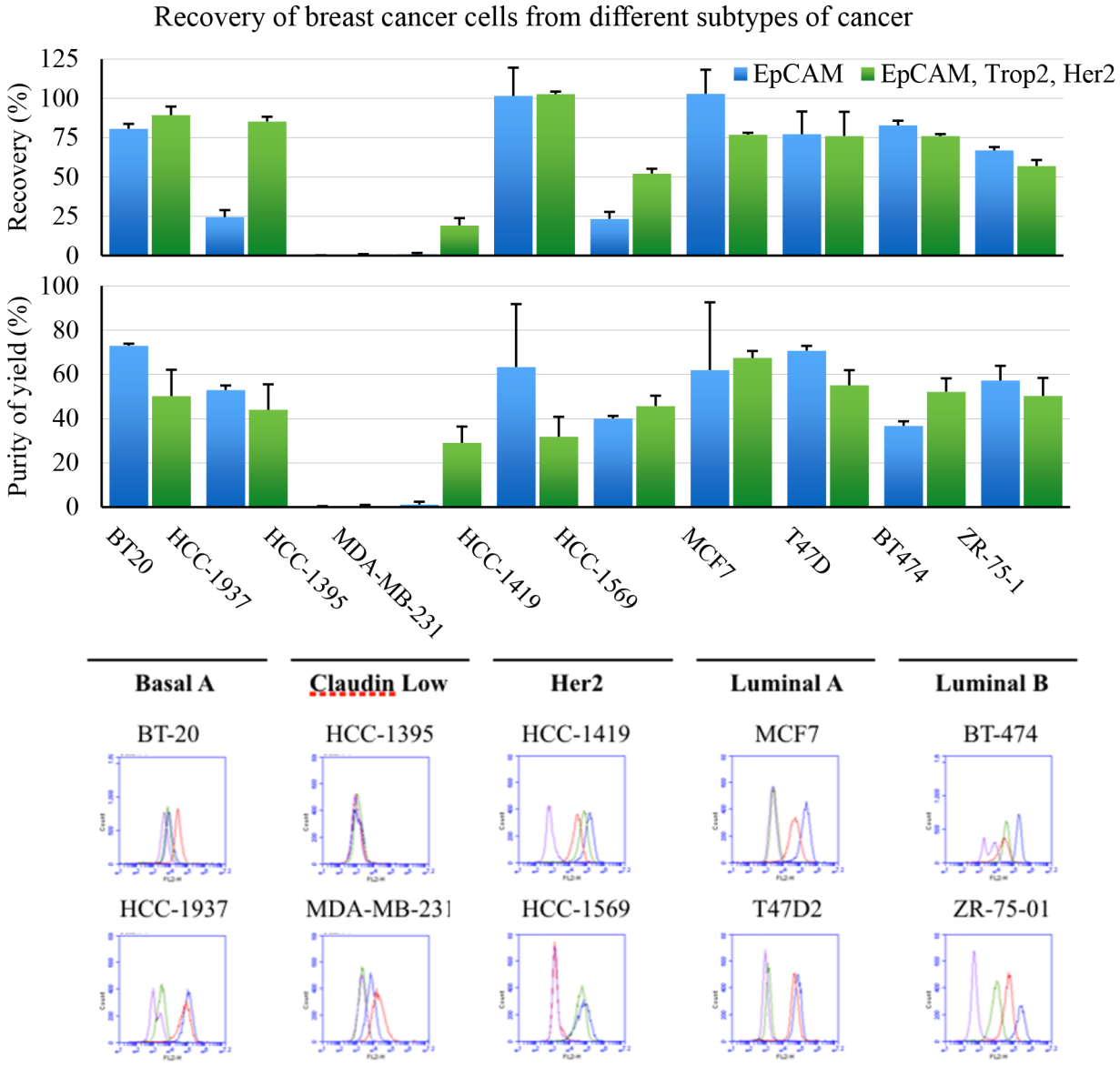

Figure 1: Tumor cells representing the five major subtypes of breast cancer were recovered from 7.5mL blood samples using the LiquidBiopsy system. The efficiency of recovery using EpCAM compared to capture with EpCAM, Her2, and Trop2 (top graph) and purity (bottom graph) are shown. Each sample was evaluated in quadruplicate and the bars represent means (+ 1 SD) The same cell lines were evaluated by FACs for control IgG1 (purple), EpCAM (blue), Trop2 (red) or Her2 expression (green) as detected using Streptavidin-FITC.

enrichment protocol was applied to 32 serial patient samples from metastatic breast cancer who were otherwise not selected for elevated CTC numbers. EpCAM based enrichment served as a control and was compared to cell populations selected with EMT cocktail. 32 blood samples were processed for EpCAM only with an average recovered purity of 7.7% Cytokeratin positive (CK+) cells. By comparison, EMT selection enriched populations with on average 8.8% cytokeratin positive cell populations with a larger range of cells than EpCAM selected (see Table 2). Importantly, the median number of CK+ cells recovered from EMT capture was almost three times elevated over EpCAM alone. (median of 23.5 cells/7.5mL vs 8.0 cells/7.5mL for EMT and EpCAM respectively). Also, the median number of CD45+ non-target cells captured was 116 and 172 cells respectively. This background is relevant to a sequencing test that supports detection of mutations present at >1%. ccfDNA sample was purified from the same tube of blood. The average concentration of ccfDNA recovered from fixed plasma samples was 7.3 ng/mL.

Evaluation of template quality

Interpreting NGS data rests on a clear understanding of the quality of template that is being interrogated. A sequencing pipeline was designed supporting case control detection of single point variants to 1% (Supplementary Figure S1). A primary distinguishing feature between ccfDNA and ctcDNA is the fragmented nature of the ccfDNA that has been described and observed previously [29, 30]. For the purposes of this experiment, the variance of amplification efficiency manifest in the ccfDNA demonstrated that approximately half of the amplicons in ccfDNA amplified the target sequence less efficiently than control (Supplementary Figure S2). To ensure sufficient coverage of the fragmented template, we therefore decided to devote one 318 chip to the analysis of each ccfDNA sample.

A second characteristic distinguishing ccfDNA and cell based DNA sequence came from the evaluation of the SNV substitution frequency (SNV-SF). The SNV-SF results in a quality metric that can be used to evaluate accuracy and precision as well as process-associated or biologically-associated noise and was used as a general metric for mutation frequency. To evaluate SNV-SF, template from the different compartments recovered from a single sample were evaluated using the AmpliSeq hotspot panel. After alignment, the template specific libraries were compared to germline sequence and only alterations with >2000 reads were evaluated. However, the reportable range for SNV-SF determination was not restricted to the Cosmic identified SNV in the Hotspot library. Therefore, any alteration detected that could not be eliminated by comparison to germline was included. This general calculation gives a comparison of different sample types from the same patient.

The SNV-SF was calculated on ccfDNA and circulating cell DNA recovered from 29 normal healthy volunteers. This analysis revealed a three-fold difference in the number of alterations observed in a known negative sample set (Figure 2). Thus, there is a noise variable manifest in ccfDNA as altered sequence that is based on the compartment of the DNA template rather than the sequencing reaction or the sequencing platform. A two-tailed Wilcoxon rank sum test measured a p value of 0.0009. If we assume that all SNVs detected in healthy donors are false positives then there is a 3-fold higher noise in the ccfDNA samples compared to the matched circulating cell DNA sample. This difference is detected on template fixed in the same sample tube, for the same time and measured with the same test. It therefore suggests the noise variable in ccfDNA is a biological noise rather than a system or assay noise. As a result, the primary comparison of ccfDNA and ctcDNA was restricted to the panel of COSMIC validated SNVs mapping to the AmpliSeq Hotspot panel in the CLIA validated test. These restrictions allowed *de novo* identification of template associated mutations across >2500 different mutations and therefore represents a toolset and workflow that supports mutation discovery on multiple templates from a single blood draw.

Clinical sequence output: Primary comparison of different templates

An important measure for clinical relevance of a NGS test is the detection of disease associated alterations in templates derived from tumor sample. This experiment was conducted on a cohort of metastatic breast cancer samples. After assembly, all variants were filtered to yield COSMIC validated mutations. For cell enrichment, we initially included EpCAM based recovery to compare capture to a cocktail of EpCAM/Her2/Trop2. As shown in Supplementary Table S1, the frequency with which EpCAM capture alone supported identification of mutation bearing cells was 9%. The EMT cocktail outperformed the EpCAM only capture by 3-fold (Table 3). Therefore, these data focus on characterizing the EMT performance.

The ctcDNA and ccfDNA samples were analyzed using a case-control model with a limit of detection of 1%. For FFPE, no case-control was used and the limit of detection was 10%. The incidence for detection of variants in evaluable FFPE samples using the AmpliSeq panel was 54% (14 of 26 with 6 samples QNS). The genes most frequently mutated were *TP53* and *PIK3CA* with mutation frequencies of 20 and 28% respectively. This is consistent with these being the most frequently altered genes in breast cancer. The frequency of mutations observed in ccfDNA and EMT ctcDNA samples was 48% and 25% respectively. Similar to the FFPE analysis, the most frequently altered genes were again *TP53* and

PIK3CA. Specifically, mutations in *TP53* and *PIK3CA* were detected with 16% and 9% frequency respectively in the ctcDNA from 32 evaluated samples. Mutations in the same genes were observed with 29% and 16% frequency in ccfDNA. In combination, ccfDNA and EMT ctcDNA produced SNV information 56% of the time - a frequency directly comparable to the FFPE sample frequency of 58%. Therefore overall, the peripheral multitemplate analysis produces signal with 98% of the frequency of evaluable FFPE samples.

The impact of sampling on heterogeneity can be observed both within a sample type as between compartments. For instance, sample CYN-026 described a *TP53* (C182Y) mutation in a bone marrow derived biopsy sample. A synchronous bone marrow sample displayed a distinct *TP53* alteration (G108S) as well as alterations in *APC* (G1447*) and *PIK3CA* (H1047R). The *PIK3CA* alteration but neither of the *TP53* or the *APC* alterations were detected in the peripheral samples. Furthermore, the *PIK3CA* alteration was detected in the EMT population

Table 2: Target cell recovery performance

	CK+ cells/7.5mL		CK+ purity (%CK+)		CD45+/ DAPI+/7.5mL
Capture Cocktail	**Median**	**sd**	**Average**	**Range**	**Median**
EpCAM	8.0	20.3	7.7%	1 - 112	116
EpCAM/Her2/Trop2	23.5	88.4	8.8%	2 - 487	172

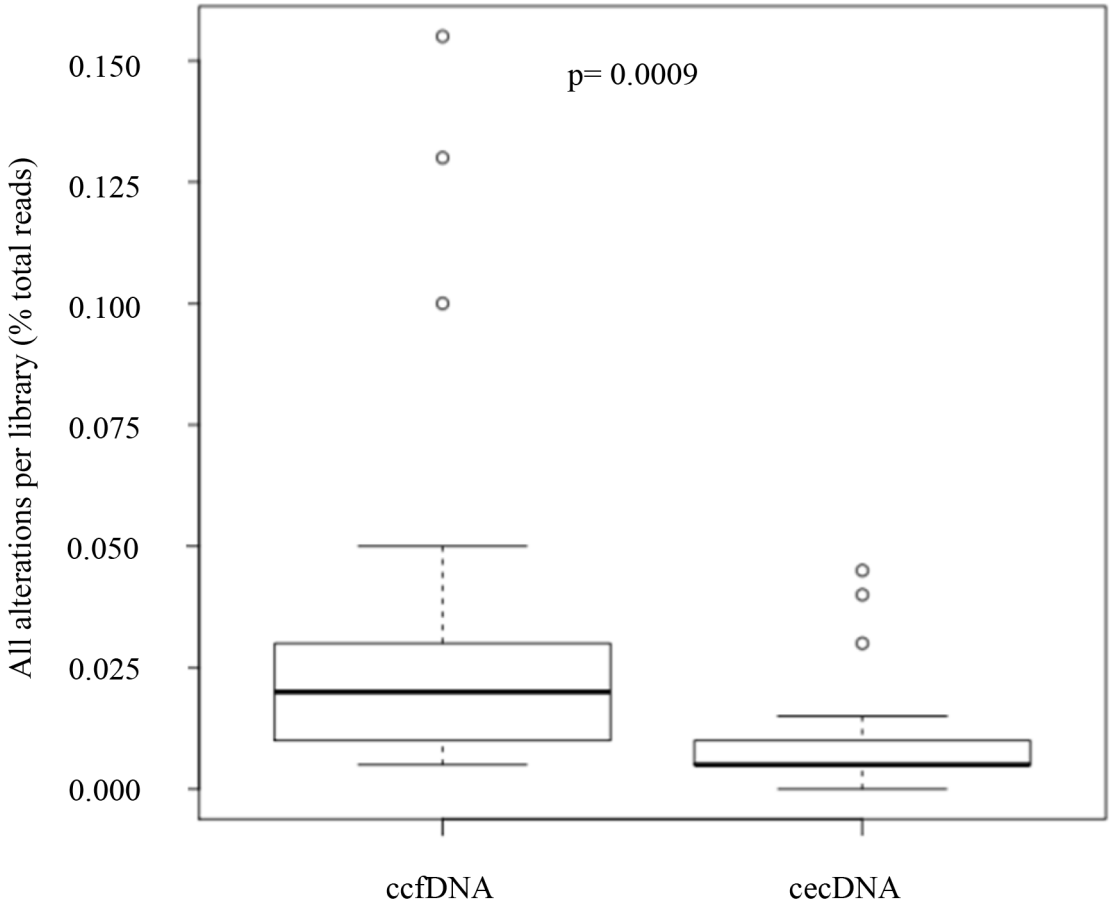

Figure 2: Evaluation of template noise in cell free DNA and cells enriched from blood as measured by SNV-SF Circulating epithelial cell DNA (cecDNA) or matched cell free DNA was recovered from 29 normal healthy donors. Paired template were sequenced using the AmpliSeq library and compared to germline sequence from the same sample. Alterations with an extended reportable range to include all evaluated bases in the library were enumerated. A two-tailed Wilcoxon rank sum test measured a mean difference of 3.0 and a p value of 0.0009.

Table 3: COSMIC identified SNV from matched tumor samples derived from blood or biopsy

ID	E/P/H	ctcDNA	ccfDNA	FFPE 1	FFPE 2
C293-001	+/+/-	X	X	X	
C293-002	-/-/+	TP53; p.E285K	X	QNS tissue	
C293-003	-/-/+	X	X	X	
C293-004	+/-/-	X	PIK3CA;p.E542K	PIK3CA;p.E542K	PIK3CA;p.E542K
C293-005	+/-/+	X	PIK3CA; p.V344G	PIK3CA; p.V344G	
			ERBB2; p.V777L	ERBB2; p.V777L	
C293-006	-/-/-	X	TP53;p.H193R	TP53;p.H193R	
C293-007	-/-/-	X	X	KRAS; V14I	
C293-008	NA	X	TP53; p.S215R	QNS tissue	
C293-009	+/+/Eq	X	X	X	
C293-010	+/-/+	TP53; p.Y163D	TP53; p.Y163D	QNS Tissue	
C293-011	-/-/-	X	TP53; p.R175H	TP53; p.R175H	
C293-012	+/-/-	X	X	X	
C293-013	+/+/-	X	QNS DNA	PIK3CA; p.H1047R	X
C293-014	+/+/Eq	X	TP53; p.M246I	X	
C293-015	-/-/-	X	X	X	
C293-016	-/-/+	PIK3CA; p.E545K	X	QNS Tissue	
C293-018	+/-/-	PIK3CA; p.H1047R	PIK3CA; p.H1047R	QNS DNA	PIK3CA; p.H1047R
C293-019	+/+/-	X	X	IDH2; p.R140Q	
C293-020	-/-/-	X	X	X	
C293-021	+/+/-	X	GNAQ; Q209K	X	
C293-022	+/+/-	X	X	PIK3CA; p.E545K	
C293-023	-/+/Eq	TP53; p.R175H	TP53; p.R175H	TP53; p.R175H	
C293-024	+/-/-	X	PIK3CA; p.H1047R	X	PIK3CA; p.H1047R
C293-025	+/-/-	X	X	X	
C293-026	+/-/Eq	PIK3CA; p.H1047R	PIK3CA; p.H1047R	TP53; p.C182Y	APC; p.Q1447*
					TP53; p.G108S
					PIK3CA; p.H1047R
C293-027	+/+/-	TP53; p.H178P	X	X	
C293-028	+/-/-	X	X	QNS DNA	
C293-029	-/-/+	X	TP53; p.E286K	QNS DNA	
C293-030	+/+/-	X	X	ERBB2; V777L	
C293-031	+/+/-	X	X	X	

(*Continued*)

ID	E/P/H	ctcDNA	ccfDNA	FFPE 1	FFPE 2
C293-032	-/-/-	X	TP53; p.Y107*	X	
C293-033	+/+/-	TP53; p.C176F	TP53; p.C176F	TP53; p.C176F	
			TP53; p.C176S		

Most recent hormonal (ER/PR) and Her2 status is included where available (E/P/H) and indicated as present (+), absent (-), or equivocal (Eq). Alterations are indicated by the target gene and the predicted impact. Samples with no detectable alteration at the limit of detection are indicated by X. Samples that were quantity not sufficient (QNS) are indicated.

but not the epithelial population. In another sample (CYN-003) a mutation in *PIK3CA* (E542K) is detected in two biopsy samples and the ccfDNA compartment but is not detected in either the epithelial or EMT cells. In contrast, CYN-016 identifies a *PIK3KCA* (E545K) driver mutation in EMT cells that is not detected in ccfDNA. The biopsy sample for the latter subject had insufficient tumor tissue on pathology review for sequence analysis. Thus, whether sampling different biopsy sites or different peripheral compartments, the three different templates are complementary.

Evaluating the impact of different tumor sampling mechanisms has been challenging due to limited studies capable of making direct comparisons between samples and indexing the template quality and performance on different templates but with the same assay. Due to the impact of insufficient biopsy material (QNS), mutations were described in 47% of the FFPE samples (26 evaluable samples of 32 subjects). Of that number, 9 samples (28%) produced peripheral signal from either ccfDNA or ctcDNA or both that was concordant with FFPE. Separately, while 72% of the samples have some information in one compartment or the other, 44% (10 of 23 samples) of any alteration could be confirmed in an orthogonal template (Table 3).

The relationship between the different sampling mechanisms is described in Figure 3. The orthogonal nature of the three sampling mechanisms detected is impacted both by the different source of sample as well as the chance that no tumor derived material is detected in either FFPE or the peripheral samples. The overlap of SNV mutation calls between sample types (FFPE, ctcDNA, ccfDNA) was compared. Concordant results were measured as the presence of any specific mutation between different sample types in the same patient. There were clinically relevant and potentially actionable mutations detected in EMT ctcDNA samples or ccfDNA samples that were not shared with other sample types. Thus EMT ccfDNA and ccfDNA produce complementary data that can supplement rare signals in either compartment or can serve as orthogonal confirmatory templates. Furthermore, the ability to tune the informative cell population can dramatically alter the frequency of informative reads.

DISCUSSION

In this prospective study, we compared NGS data sets from three distinct patient-matched samples types (FFPE, ctcDNA, ccfDNA). A common amplicon based resequencing panel (Ampliseq v2 HotSpot panel) was used for all template types in a NGS pipeline using identical variant analysis. Patients with available quality tissue specimen from diagnostic biopsy were isolated and compared. We demonstrated that ccfDNA and ctcDNA evaluation yields complementary molecular information from the same blood sample. Moreover, the capture of CTC for molecular analysis may be tuned based on a molecular detection system. In fact, the capture protocol for circulating mutation bearing cells was a novel cocktail of anti-EpCAM, Her2, and Trop2 antibodies. This cocktail was designed based on the established role of EpCAM in defining a population of cells in blood that is prognostically related to disease progression as well as metastatic events [31, 32]. The additional markers have established an emerging utility for antibody directed therapies in breast cancer targeting Her2 and Trop2 respectively [33-36]. While other cocktails have been tested, only a subset has been shown to be productive for the detection of mutation bearing cells [11, 37]. Furthermore, in our study the detection and molecular analysis of CTCs was not dependent on clinical or biological outcome to interpret the value of the populations. Specifically, the CellSearch test is not dependent only on EpCAM expression *per se*. Rather, if anything changes, including cells or reagents, the approved prognostic value is lost. For these applications, the most important result is the demonstration that the informative template from a blood sample may now be defined molecularly as related to the disease process by orthogonal characterization of DNA alterations. DNA based analysis of SNV was utilized to demonstrate qualitative characteristics between the different template compartments. Having validated the capture definition, the informative cell population may now be interrogated for additional biomarkers of value such as expression of protein markers [38], protein modification events, or RNA expression based analysis for biomarkers that are opaque in the DNA compartment such as the AR-V7 splice variant in prostate cancer [39].

The additional goal of these comparisons was to evaluate different template compartments for the ability to provide clinically useful DNA sequence information. The measures of success for these comparisons were deemed to be i) the frequency of disease associated alterations observed in the analysis of multiple templates (tissue, ctcDNA or ccfDNA) and ii) the quality of template for supporting sequence analysis. A key observation is that the FFPE sample and the complementary peripheral samples together provide molecular information in a similar proportion of samples. In these data, all templates are detecting alterations within a two-fold range of each other. Previous studies comparing CTC and ccfDNA have been limited by making comparisons with different technologies [21] or using shared readouts that do not alter the significant signal to noise challenge of detecting rare mutation bearing cells [23]. These approaches made it challenging to perform a direct comparison of ccfDNA and ctcDNA. A recent report demonstrating the presence of a T790M resistance mutation in lung cancer supports the data presented here in that ctcDNA and ccfDNA produce complementary information [26].

Fundamental to this experiment is the ability to clarify the population of cells molecularly. There are numerous approaches to recovering tumor cells from blood samples ranging from size based selection, to mechanical mechanisms, charge based approaches or no selection. Each has applications for discovery of previously uncharacterized cell populations. In a patient setting, population based (as opposed to single cell) analysis of cells or ccfDNA fragments allows more rapid molecular analysis but also uses a sampling mechanism that may be less sensitive to analytic variables. A significant advantage of positive selection approaches is the ability to standardize the enrichment definition around a standardizable biological definition.

The studies reported here are based upon the examination of cell pools, as well as mixtures of ccfDNA isolated from whole blood. To the best of our knowledge, there are no published data that show a population of tumor cells in blood can be productively recovered and sequenced directly yielding a clinically relevant mutational spectrum. Other direct evidence for the significant presence of mutation bearing cells has come from single cell experiments or single cell

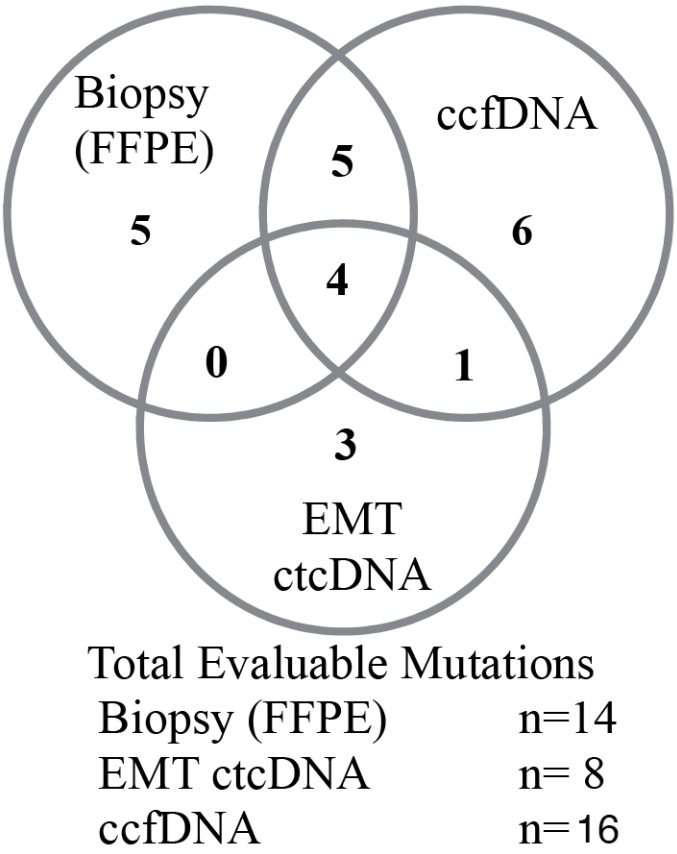

Figure 3: COSMIC identified SNV from matched tumor samples derived from blood or biopsy. Alterations are counted in all clinical samples. Overlap between each template is depicted by Venn diagram with comparison between FFPE, ccfDNA and ctcDNA compartments.

based readouts such as FISH analysis [10, 40-42]. Despite that, a number of studies demonstrated the presence of cells in blood that contain mutations that can be tracked back to the tumor but almost all are restricted single cell selection [43], single cell analysis [44, 45], or single cell propagation steps [11, 37]. The population based approach we present in this paper supports high throughput based analysis. In addition, a median background of 172 CD45+ events from a 7.5mL blood sample will allow detection of as few as 3-4 heterozygous events from a blood sample using a test that can detect mutations with a 1% frequency; a performance already approaching a single cell detection event. Furthermore, individuals without cancer exhibit recoverable cells with a phenotype related to circulating epithelial cells [46, 47]. However these cells do not bear cancerous SNV mutations. It will be important therefore to develop advanced validation tools, like detection of DNA alterations, before further characterization of cell populations with tools such as expression profiling. Even enumeration has to be indexed to a clinically prognostic impact or molecular demonstration that the cell population is truly tumor derived.

The current data demonstrate that both blood derived templates support informative amplicon based resequencing. However, significant differences between the templates with respect to their performance in NGS variant analysis suggest that the templates need to be handled according to their strengths and weaknesses by standardizing the analysis rules. Key among these observations was the demonstration of a variant noise detected in the ccfDNA that is significantly elevated over cell based analysis from the same sample. This noise was detected at random sites within the amplicon that are not associated with the disease process and need to be eliminated by a restriction call using bioinformatic methods. This has implications for the ability to use this approach to map *de novo* mutations discovered in the ccfDNA template. One hypothesis to explain this noise, postulates that noise is due to the fixative or preservative utilized, analogous to the template damage observed with FFPE template [48]. However the absence of a similar noise profile in other matched DNA templates recovered from the same sample argues against this hypothesis. ccfDNA is a degradative product of catabolism, present in the extracellular compartment and therefore is susceptible to multiple processes of damage in plasma. Indeed, when specifically evaluated, a similar background has been described elsewhere [49]. Alternatively, given the data suggesting accumulation of DNA damage over time in preneoplastic tissues, we can speculate that this template associated noise might be a reflection of that larger phenomenon in biology, in essence detecting a level of "normal" DNA damage even in the absence of diagnosed malignancies with potential implications for future studies [50-52]. Certainly, this difference between ctcDNA and ccfDNA quality reflects the known differences in their biological source and has implications for the prognostic value of each biomarker [53].

Clearly, detection of tumor related events depend on the targets included in the NGS test. The AmpliSeq v2 Hotspot panel is a well curated pan-cancer panel that has been shown to identify disease associated alterations in FFPE and ccfDNA [54-56]. The most frequently mutated alterations in breast cancer are detected with related frequencies in the multiple templates. However, not all disease associated mutations can be included in any targeted panel. More comprehensive whole genome or whole exome sequencing is possible but should be considered in the context of the enzymatic amplification steps necessary to produce sufficient template from limiting numbers of cells and the impact on sequence integrity. Therefore, it will be key to develop detection tools that reflect the most relevant decision points for any given patient population. For instance, in breast cancer, there are clinically relevant biomarker targets in the protein, RNA or DNA compartments [57-63]. Functional access to both ccfDNA and CTC based templates supports development of protein and RNA based readouts in addition to detection of DNA based alterations.

This study demonstrates that it is possible to productively interrogate both populations of mutation bearing cells enriched from a blood sample and ccfDNA. Importantly, this analysis does not require *a priori* knowledge of the mutations present and will serve as a useful discovery tool. Expanding the definition of informative cell populations may now be indexed to a molecular definition. This ability to define populations beyond the legacy epithelial definition raises the possibility of informative biomarkers in non-epithelial settings such as soft tissue cancers, skin cancers such as melanoma, or epithelial cancers that have not previously performed well with strictly epithelial markers but still characterized by metastatic events.

Fundamentally, tumor derived samples from biopsy, circulating tumor cell populations and ccfDNA are acquired by different sampling mechanisms; mechanisms that capture different moments of the disease and potentially not necessarily the same biological sources. The study of these alternate templates by next generation sequencing technologies has emphasized the clonal nature of cancer and the impact of evaluating these different biopsy sites [64]. The emerging model for the peripheral templates suggest ccfDNA samples the genomic DNA fragments released from all tumor sites. In contrast, circulating tumor cells are clearly related to the disease process, predict more aggressive disease and increased metastasis [65]. As such, CTC reflect the mobile and

metastatic subset of tumor cells. Clearly the overlap between detectable alterations in each compartment demonstrates the relevant information may not always be linked to one specific compartment or one biology. An ability to evaluate multiple compartments, both in terms of sensitivity and biomarker definition, is an innovative approach that can have significant impact with regards to diagnosis and on our capacity to better understand the various biological processes driving metastases and potentially different therapeutic approaches. The impact of these biological differences will have to be determined in prospective studies.

MATERIALS AND METHODS

Ethics statement

This prospective clinical trial was funded with a NIH/NCI contract with the express purpose of expanding the definition of evaluable tumor cells in blood. The long term goal of the SBIR contract was to develop new devices and methods of CTC detection with a focus on clinical tools rather than research based approaches. Patients were enrolled at three different sites (Kimmel Cancer Center of Thomas Jefferson University, Winship Cancer Institute of Emory University, and The West Clinic). Normal human donor blood was purchased from HemaCare Corporation. All subjects provided written informed consent under an Institutional Review Board approved protocols.

Study subjects

The study was designed to enroll subjects with a confirmed diagnosis of metastatic breast cancer that were about to start a new line of therapy for their disease. A tissue biopsy of the metastatic lesion was mandatory. After appropriate consent enrolled subjects provided a blood sample before starting the new therapy. The blood samples had a minimum volume of 16mL minimum from two K_2-EDTA tubes with minimal signs of hemolysis. Samples were processed for recovery of CTC populations within 96 hours using a CLIA approved process. Thirty-two metastatic breast cancer patients were recruited. FFPE and blood was recovered from each donor. From the blood samples, germline white blood cell DNA (wbcDNA), ccfDNA, and ctcDNA were isolated. Sequencing libraries were constructed from the matched samples and subjected to NGS using the IonTorrent PGM platform.

Template enrichment procedures

All sample processing, sequencing and analysis was performed in the Cynvenio Biosystems CAP approved facility (Westlake Village, CA) under CLIA supervision. Whole blood was collected in purple top (K_2EDTA) tubes and stabilized using LiquidBiopsy (Cynvenio Biosystems) fixative. A white blood cell control was recovered from 0.1cc of the original sample. Plasma was collected after brief centrifugation to separate cellular components. CTC's were enriched as described [25]. In brief, the cellular component in the starting blood volume was blocked with FcR block and labelled with a biotinylated antibody cocktail consisting of anti-EpCAM alone, or in combination with anti-HER2, and anti-TROP2 (Cynvenio Biosystems) followed by iMAG streptavidin beads. The labeled blood was processed in the CTC flow cell on the LiquidBiopsy platform (Cynvenio Biosystems). Captured cells were characterized by evaluating immunofluorescent staining with anti-Cytokeratin, anti-CD45 and DAPI. Captured cells were recovered by centrifugation to produce an enriched cell pellet. The CTC pellet was digested as described [25] and the resulting digest was diluted to 12µL with TE. The AmpliSeq library reagents were added directly to the template and further processed as for the germline and ccfDNA samples.

For enrichment of ccfDNA from the recovered plasma, the QIAamp Circulating Nucleic Acid Kit was used along with the QIAVAC system as recommended by the manufacturer (Qiagen, Valencia, CA). DNA resulting from this purification was quantitated on a Nanodrop (Thermo Fisher Scientific) and directly utilized for sequencing library generation. ccfDNA libraries were produced with a 10 ng input of ccfDNA.

FFPE processed slides were H&E stained, graded by pathologist to indicate the area of specific tumor tissue, macro-dissected, and placed into individual tubes. Three to six 5µm thick sections were processed using the Agencourt® FormaPure® kit (Beckman Coulter). Samples were eluted in 45µL and DNA was concentrated further with Agencourt AMPure XP beads to yield a final DNA volume of 15µL. DNA concentration was measured by Qubit® dsDNA HS (High Sensitivity) Assay Kit (Thermo Fisher Scientific).

Sequencing data analysis

Primary sequence was demultiplexed and exported from the Torrent Server as FASTQ. The FASTQ files were aligned by reference guided assembly to NCBI GRCh37 p5 using Bowtie 2 [66]. Post-assembly alignments were piled and curated for accuracy using SAM Tools (version 0.1.19) [67] and transferred to Perl using Bio::DB::Sam. ctcDNA and ccfDNA templates were analyzed and a mutation was called if ≥20 mutant reads were observed for a limit of detection (LOD) of 1%. FFPE templates were analyzed to a LOD of 10%. ctcDNA and ccfDNA analysis was based upon a case-control model for variant detection, in which total read coverage must be ≥ 2000 reads per amplicon for ctcDNA and ccfDNA sample and validated calls were required to be absent from the negative control wbcDNA sequence (Supplementary Figure S1). FFPE analysis was not case-controlled, and total read coverage threshold was ≥ 500 reads per amplicon.

ACKNOWLEDGMENTS

We would like to thank the staff and patients who took part in this study and made it possible.

CONFLICTS OF INTEREST

WMS, CC, JS, EK, BV, JR, MMS, and PWD are employees of Cynvenio Biosystems. WMS, BV, and PWD are inventors on issued patents or patent applications associated with this work and owned by Cynvenio Biosystems.

GRANT SUPPORT

The clinical study was supported by the Small Business Innovation Research Program (SBIR), RFP No. N44C031014-51 SBIR Phase II Topic 293: *Development of Devices for Point of Care Analysis of Circulating Tumor Cells.* Contract No. HHSN261201300073C.

REFERENCES

1. Hansen AR, Bedard PL. Clinical application of high-throughput genomic technologies for treatment selection in breast cancer. Breast Cancer Res. 2013; 15:R97.
2. Gerlinger M, Rowan AJ, Horswell S, Larkin J, Endesfelder D, Gronroos E, Martinez P, Matthews N, Stewart A, Tarpey P, Varela I, Phillimore B, Begum S, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. N Engl J Med. 2012; 366:883–892.
3. Yap TA, Gerlinger M, Futreal PA, Pusztai L, Swanton C. Intratumor heterogeneity: seeing the wood for the trees. Sci Transl Med. American Association for the Advancement of Science. 2012; 4:127ps10–0.
4. De Mattos-Arruda L, Cortes J, Santarpia L, Vivancos A, Tabernero J, Reis-Filho JS, Seoane J. Circulating tumour cells and cell-free DNA as tools for managing breast cancer. Nat Rev Clin Oncol. 2013; 10:377–389.
5. Ring AE, Zabaglo L, Ormerod MG, Smith IE, Dowsett M. Detection of circulating epithelial cells in the blood of patients with breast cancer: comparison of three techniques. British Journal of Cancer. 2005; 92:906–912.
6. Racila E, Euhus D, Weiss AJ, Rao C, McConnell J, Terstappen LW, Uhr JW. Detection and characterization of carcinoma cells in the blood. Proc Natl Acad Sci USA. 1998; 95:4589–4594.
7. Cristofanilli M, Budd GT, Ellis MJ, Stopeck A, Matera J, Miller MC, Reuben JM, Doyle GV, Allard WJ, Terstappen LWMM, Hayes DF. Circulating tumor cells, disease progression, and survival in metastatic breast cancer. N Engl J Med. 2004; 351:781–791.
8. Fernandez SV, Bingham C, Fittipaldi P, Austin L, Palazzo J, Palmer G, Alpaugh K, Cristofanilli M. TP53 mutations detected in circulating tumor cells present in the blood of metastatic triple negative breast cancer patients. Breast Cancer Res. 2014; 16:445-456.
9. Mayer JA, Pham T, Wong KL, Scoggin J, Sales EV, Clarin T, Pircher TJ, Mikolajczyk SD, Cotter PD, Bischoff FZ. FISH-based determination of HER2 status in circulating tumor cells isolated with the microfluidic CEE platform. Cancer Genetics. 2011; 204:589–595.
10. Punnoose EA, Atwal SK, Spoerke JM, Savage H, Pandita A, Yeh R-F, Pirzkall A, Fine BM, Amler LC, Chen DS, Lackner MR. Molecular biomarker analyses using circulating tumor cells. PLoS ONE. 2010; 5:e12517.
11. Yu M, Bardia A, Wittner BS, Stott SL, Smas ME, Ting DT, Isakoff SJ, Ciciliano JC, Wells MN, Shah AM, Concannon KF, Donaldson MC, Sequist LV, et al. Circulating breast tumor cells exhibit dynamic changes in epithelial and mesenchymal composition. Science. 2013; 339:580–584.
12. Mego M, Gao H, Lee B-N, Cohen EN, Tin S, Giordano A, Wu Q, Liu P, Nieto Y, Champlin RE, Hortobagyi GN, Cristofanilli M, Ueno NT, et al. Prognostic Value of EMT-Circulating Tumor Cells in Metastatic Breast Cancer Patients Undergoing High-Dose Chemotherapy with Autologous Hematopoietic Stem Cell Transplantation. J Cancer. 2012; 3:369–380.
13. Joosse SA, Pantel K. Biologic challenges in the detection of circulating tumor cells. Cancer Res. 2013; 73:8–11.
14. Joosse SA, Hannermann J, Spötter J, Bauche A, Andreas A, Müller V, Pantel K. Changes in Keratin Expression during Metastatic Progression of Breast Cancer: Impact on the Detection of Circulating Tumor Cells. Clinical Cancer Research. 2013; 18:993–1003.
15. Smerage JB, Barlow WE, Hortobagyi GN, Winer EP, Leyland-Jones B, Srkalovic G, Tejwani S, Schott AF, O'Rourke MA, Lew DL, Doyle GV, Gralow JR, Livingston RB, et al. Circulating Tumor Cells and Response to Chemotherapy in Metastatic Breast Cancer: SWOG S0500. Journal of Clinical Oncology. 2014; 32:3483–3489.
16. Heitzer E, Auer M, Gasch C, Pichler M, Ulz P, Hoffmann EM, Lax S, Waldispuehl-Geigl J, Mauermann O, Lackner C, Hofler G, Eisner F, Sill H, et al. Complex Tumor Genomes Inferred from Single Circulating Tumor Cells by Array-CGH and Next-Generation Sequencing. Cancer Res. 2013; 73:2965–2975.
17. Danila DC, Anand A, Sung CC, Heller G, Leversha MA, Cao L, Lilja H, Molina A, Sawyers CL, Fleisher M, Scher HI. TMPRSS2-ERG status in circulating tumor cells as a predictive biomarker of sensitivity in castration-resistant prostate cancer patients treated with abiraterone acetate. Eur Urol. 2011; 60:897–904.
18. Mandel P. Les acides nucleiques du plasma sanguin chez l'homme. CR Acad Sci Paris. 1948; 142:241–243.
19. Diehl F, Schmidt K, Choti MA, Romans K, Goodman S, Li M, Thornton K, Agrawal N, Sokoll L, Szabo SA, Kinzler

KW, Vogelstein B, Diaz LA. Circulating mutant DNA to assess tumor dynamics. Nat Med. 2008; 14:985–990.

20. Murtaza M, Dawson S-J, Tsui DWY, Gale D, Forshew T, Psikorz AM, Parkinson C, Chin S-F, Kingsbury Z, Wong ASC, Marass F, Humphrey S, Hadfield J, et al. Non-invasive analysis of acquired resistance to cancer therapy by sequencing of plasma DNA. Nature. 2013; 497:108–112.
21. Dawson S-J, Tsui DWY, Murtaza M, Biggs H, Rueda OM, Chin S-F, Dunning MJ, Gale D, Forshew T, Mahler-Araujo B, Rajan S, Humphray S, Becq J, Halsall D, Wallis M, Bentley D, Caldas C, Rosenfeld N. Analysis of Circulating Tumor DNA to Monitor Metastatic Breast Cancer. N Engl J Med. 2013; 368:1199-1209.
22. Forshew T, Murtaza M, Parkinson C, Gale D, Tsui DWY, Kaper F, Dawson S-J, Piskorz AM, Jimenez-Linan M, Bentley D, Hadfield J, May AP, Caldas C, et al. Noninvasive identification and monitoring of cancer mutations by targeted deep sequencing of plasma DNA. Sci Transl Med. 2012; 4:136ra68.
23. Bettegowda C, Sausen M, Leary RJ, Kinde I, Wang Y, Agrawal N, Bartlett BR, Wang H, Luber B, Alani RM, Antonarakis ES, Azad NS, Bardelli A, et al. Detection of circulating tumor DNA in early- and late-stage human malignancies. Sci Transl Med. 2014; 6:224ra24–4.
24. Jahr S, Hentze H, Englisch S, Hardt D, Fackelmayer FO, Hesch R-D, Knippers R. DNA Fragments in the Blood Plasma of Cancer Patients: Quantitations and Evidence for Their Origin from Apoptotic and Necrotic Cells. Cancer Res. 2001; 61:1659–1665.
25. Thierry AR, Mouliere F, Gongora C, Ollier J, Robert B, Ychou M, Del Rio M, Molina F. Origin and quantification of circulating DNA in mice with human colorectal cancer xenografts. Nucleic Acids Research. 2010; 38:6159–6175.
26. Sundaresan TK, Sequist LV, Heymach JV, Riely GJ, Jänne PA, Koch WH, Sullivan JP, Fox DB, Maher R, Muzikansky A, Webb A, Tran HT, Giri U, et al. Detection of T790M, the acquired resistance EGFR mutation, by tumor biopsy versus noninvasive blood-based analyses. Clin Cancer Res. 2015; :1–32.
27. Winer-Jones JP, Vahidi B, Arquilevich N, Fang C, Ferguson S, Harkins D, Hill C, Klem E, Pagano PC, Peasley C, Romero J, Shartle R, Vasko RC, et al. Circulating Tumor Cells: Clinically Relevant Molecular Access Based on a Novel CTC Flow Cell. PLoS ONE. 2014; 9:e86717.
28. Prat A, Karginova O, Parker JS, Fan C, He X, Bixby L, Harrell JC, Roman E, Adamo B, Troester M, Perou CM. Characterization of cell lines derived from breast cancers and normal mammary tissues for the study of the intrinsic molecular subtypes. Breast Cancer Res Treat. 2013; 142:237–255.
29. Devonshire AS, Whale AS, Gutteridge A, Jones G, Cowen S, Foy CA, Huggett JF. Towards standardisation of cell-free DNA measurement in plasma: controls for extraction efficiency, fragment size bias and quantification. Anal Bioanal Chem. 2014; 406:6499-6512.
30. Wang BG, Huang H-Y, Chen Y-C, Bristow RE, Kassauei K, Cheng C-C, Roden R, Sokoll LJ, Chan DW, Shih I-M. Increased Plasma DNA Integrity in Cancer Patients. Cancer Res. 2003; 63:3966–3968.
31. Lucci A, Hall CS, Lodhi AK, Bhattacharyya A, Anderson AE, Xiao L, Bedrosian I, Kuerer HM, Krishnamurthy S. Circulating tumour cells in non-metastatic breast cancer: a prospective study. Lancet Oncol. 2012; 13:688-695.
32. Rack B, Schindlbeck C, Juckstock J, Andergassen U, Hepp P, Zwingers T, Friedl TWP, Lorenz R, Tesch H, Fasching PA, Fehm T, Schneeweiss A, Lichtenegger W, et al. Circulating Tumor Cells Predict Survival in Early Average-to-High Risk Breast Cancer Patients. Journal of the National Cancer Institute. 2014; 106:dju066.
33. Liu D, Cardillo TM, Wang Y, Rossi EA, Goldenberg DM, Chang C-H. Trop-2-targeting tetrakis-ranpirnase has potent antitumor activity against triple-negative breast cancer. Mol Cancer. 2014; 13:1–12.
34. Lin H, Zhang H, Wang J, Lu M, Zheng F, Wang C, Tang X, Xu N, Chen R, Zhang D, Zhao P, Zhu J, Mao Y, et al. A novel human Fab antibody for Trop2 inhibits breast cancer growth in vitro and in vivo. Int J Cancer. 2013; 134:1239–1249.
35. Thierry B, Kurkuri M, Shi JY, Lwin LEMP, Palms D. Herceptin functionalized microfluidic polydimethylsiloxane devices for the capture of human epidermal growth factor receptor 2 positive circulating breast cancer cells. Biomicrofluidics. 2010; 4:032205.
36. Meng S, Tripathy D, Shete S, Ashfaq R, Haley B, Perkins S, Beitsch P, Khan A, Euhus D, Osborne C, Frenkel E, Hoover S, Leitch M, et al. HER-2 gene amplification can be acquired as breast cancer progresses. Proc Natl Acad Sci USA. 2004; 101:9393–9398.
37. Zhang L, Ridgway LD, Wetzel MD, Ngo J, Yin W, Kumar D, Goodman JC, Groves MD, Marchetti D. The Identification and Characterization of Breast Cancer CTCs Competent for Brain Metastasis. Sci Transl Med. 2013; 5:180ra48–8.
38. Yokobori T, Iinuma H, Shimamura T, Imoto S, Sugimachi K, Ishii H, Iwatsuki M, Ota D, Ohkuma M, Iwaya T, Nishida N, Kogo R, Sudo T, et al. Plastin3 is a novel marker for circulating tumor cells undergoing the epithelial-mesenchymal transition and is associated with colorectal cancer prognosis. Cancer Res. 2013; 73:2059–2069.
39. Antonarakis ES, Lu C, Wang H, Luber B, Nakazawa M, Roeser JC, Chen Y, Mohammad TA, Chen Y, Fedor HL, Lotan TL, Zheng Q, De Marzo AM, et al. AR-V7 and Resistance to Enzalutamide and Abiraterone in Prostate Cancer. N Engl J Med. 2014; 371:1028-1038.
40. Pecot CV, Bischoff FZ, Mayer JA, Wong KL, Pham T, Bottsford-Miller J, Stone RL, Lin YG, Jaladurgam P, Roh JW, Goodman BW, Merritt WM, Pircher TJ, et al. A novel platform for detection of CK+ and CK- CTCs. Cancer Discovery. 2011; 1:580–586.

41. Krishnamurthy S, Bischoff F, Ann Mayer J, Wong K, Pham T, Kuerer H, Lodhi A, Bhattacharyya A, Hall C, Lucci A. Discordance in HER2gene amplification in circulating and disseminated tumor cells in patients with operable breast cancer. Cancer Med. 2013; 2:226–233.

42. Gasch C, Bauernhofer T, Pichler M, Langer-Freitag S, Reeh M, Seifert AM, Mauermann O, Izbicki JR, Pantel K, Riethdorf S. Heterogeneity of Epidermal Growth Factor Receptor Status and Mutations of KRAS/PIK3CA in Circulating Tumor Cells of Patients with Colorectal Cancer. Clinical Chemistry. 2012; 59:252–260.

43. Polzer B, Medoro G, Pasch S, Fontana F, Zorzino L, Pestka A, Andergassen U, Meier-Stiegen F, Czyz ZT, Alberter B, Treitschke S, Schamberger T, Sergio M, et al. Molecular profiling of single circulating tumor cells with diagnostic intention. EMBO Mol Med. 2014; 6:1371–1386.

44. Sakaizawa K, Goto Y, Kiniwa Y, Uchiyama A, Harada K, Shimada S, Saida T, Ferrone S, Takata M, Uhara H, Okuyama R. Mutation analysis of BRAF and KIT in circulating melanoma cells at the single cell level. British Journal of Cancer. 2012; 106:939–946.

45. Klein CA, Schmidt-Kittler O, Blankenstein T, Petronio M, Polzer B, Stoecklein NH, Riethmüller G. Genetic heterogeneity of single disseminated tumour cells in minimal residual cancer. Lancet. 2002; 360:683–639.

46. Pantel K, Denève E, Nocca D, Coffy A, Vendrell J-P, Maudelonde T, Riethdorf S, Alix-Panabières C. Circulating epithelial cells in patients with benign colon diseases. Clinical Chemistry. 2012; 58:936–940.

47. Stott SL, Lee RJ, Nagrath S, Yu M, Miyamoto DT, Ulkus L, Inserra EJ, Ulman M, Springer S, Nakamura Z, Moore AL, Tsukrov DI, Kempner ME, et al. Isolation and characterization of circulating tumor cells from patients with localized and metastatic prostate cancer. Sci Transl Med. 2010; 2:25ra23.

48. Wong SQ, Li J, Tan AY-C, Vedururu R, Pang J-MB, Do H, Ellul J, Doig K, Bell A, MacArthur GA, Fox SB, Thomas DM, Fellowes A, et al. Sequence artefacts in a prospective series of formalin-fixed tumours tested for mutations in hotspot regions by massively parallel sequencing. BMC Med Genomics. 2014; 7:23-33.

49. Newman AM, Bratman SV, To J, Wynne JF, Eclov NCW, Modlin LA, Liu CL, Neal JW, Wakelee HA, Merritt RE, Shrager JB, Loo BW, Alizadeh AA, et al. An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage. Nat Med. 2014; 20:548–554.

50. Genovese G, Kähler AK, Handsaker RE, Lindberg J, Rose SA, Bakhoum SF, Chambert K, Mick E, Neale BM, Fromer M, Purcell SM, Svantesson O, Landén M, et al. Clonal Hematopoiesis and Blood-Cancer Risk Inferred from Blood DNA Sequence. N Engl J Med. 2014; 371:2477–2487.

51. Behjati S, Huch M, van Boxtel R, Karthaus W, Wedge DC, Tamuri AU, Martincorena I, Petljak M, Alexandrov LB, Gundem G, Tarpey PS, Roerink S, Blokker J, et al. Genome sequencing of normal cells reveals developmental lineages and mutational processes. Nature. 2014; 513:422–425.

52. Vijg J. Somatic mutations, genome mosaicism, cancer and aging. Curr Opin Genet Dev. 2014; 26:141–149.

53. Madic J, Kiialainen A, Bidard F-C, Birzele F, Ramey G, Leroy Q, Frio TR, Vaucher I, Raynal V, Bernard V, Lermine A, Clausen I, Giroud N, et al. Circulating tumor DNA and circulating tumor cells in metastatic triple negative breast cancer patients. Int J Cancer. 2014; 136:2158–2165.

54. Bai X, Zhang E, Ye H, Nandakumar V, Wang Z, Chen L, Tang C, Li J, Li H, Zhang W, Han W, Lou F, Zhang D, et al. PIK3CA and TP53 gene mutations in human breast cancer tumors frequently detected by ion torrent DNA sequencing. PLoS ONE. 2014; 9:e99306.

55. Tsongalis GJ, Peterson JD, de Abreu FB, Tunkey CD, Gallagher TL, Strausbaugh LD, Wells WA, Amos CI. Routine use of the Ion Torrent AmpliSeq™ Cancer Hotspot Panel for identification of clinically actionable somatic mutations. Clin Chem Lab Med. 2014; 52:707–714.

56. Rothé F, Laes JF, Lambrechts D, Smeets D, Vincent D, Maetens M, Fumagalli D, Michiels S, Drisis S, Moerman C, Detiffe JP, Larsimont D, Awada A, et al. Plasma circulating tumor DNA as an alternative to metastatic biopsies for mutational analysis in breast cancer. Annals of Oncology. 2014; 25:1959–1965.

57. Mostert B, Sieuwerts AM, Kraan J, Bolt-de Vries J, van der Spoel P, Van Galen A, Peeters DJ, Dirix LY, Seynaeve CM, Jager A, de Jongh FE, Hamberg P, Stouthard JML, et al. Gene expression profiles in circulating tumor cells to predict prognosis in metastatic breast cancer patients. Annals of Oncology. 2015; 26:510–516.

58. Kalinsky K, Mayer JA, Xu X, Pham T, Wong KL, Villarin E, Pircher TJ, Brown M, Maurer MA, Bischoff FZ. Correlation of hormone receptor status between circulating tumor cells, primary tumor, and metastasis in breast cancer patients. Clin Transl Oncol. 2015; 17:539–546.

59. Lang JE, Scott JH, Wolf DM, Novak P, Punj V, Magbanua MJM, Zhu W, Mineyev N, Haqq CM, Crothers JR, Esserman LJ, Tripathy D, van t Veer L, et al. Expression profiling of circulating tumor cells in metastatic breast cancer. Breast Cancer Res Treat. 2014; 149:121-131.

60. Andre F, Bachelot T, Commo F, Campone M, Arnedos M, Dieras V, Lacroix-Triki M, Lacroix L, Cohen P, Gentien D, Adélaide J, Dalenc F, Goncalves A, et al. Comparative genomic hybridisation array and DNA sequencing to direct treatment of metastatic breast cancer: a multicentre, prospective trial (SAFIR01/UNICANCER). Lancet Oncology. 2014; 15:267–274.

61. Banerji S, Cibulskis K, Rangel-Escareno C, Brown KK, Carter SL, Frederick AM, Lawrence MS, Sivachenko AY, Sougnez C, Zou L, Cortes ML, Fernandez-Lopez JC, Peng S, et al. Sequence analysis of mutations and translocations across breast cancer subtypes. Nature. 2012; 486:405-409.

62. Nik-Zainal S, Alexandrov LB, Wedge DC, Van Loo P, Greenman CD, Raine K, Jones D, Hinton J, Marshall J, Stebbings LA, Menzies A, Martin S, Leung K, et al. Mutational Processes Molding the Genomes of 21 Breast Cancers. Cell. 2012; 149:979–993.

63. Curtis C, Shah SP, Chin S-F, Dunning MJ, Lynch AG, Samarajiwa S, Yuan Y, Bashashati A, METABRIC Group, Langerød A, Provenzano E, Børresen-Dale A-L, Brenton JD, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature. 2012; 486:346–352.

64. Criscitiello C, Andre F, Thompson AM, De Laurentiis M, Esposito A, Gelao L, Fumagalli L, Locatelli M, Minchella I, Orsi F, Goldhirsch A, Curigliano G. Biopsy confirmation of metastatic sites in breast cancer patients: clinical impact and future perspectives. Breast Cancer Res. 2014; 16:205-216.

65. Giuliano M, Giordano A, Jackson S, De Giorgi U, Mego M, Cohen EN, Gao H, Anfossi S, Handy BC, Ueno NT, Alvarez RH, De Placido S, Valero V, et al. Circulating tumor cells as early predictors of metastatic spread in breast cancer patients with limited metastatic dissemination. Breast Cancer Res. 2014; 16:440-449.

66. Langmead B, Salzberg SL, Fast gapped-read alignment with Bowtie 2. Nat Meth. 2012; 9: 357–359.

67. Li H, A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics. 2011; 27:2987–2993.