



# Circulating Tumor Cell Transcriptomics as Biopsy Surrogates in Metastatic Breast Cancer

Alexander Ring, MD, PhD<sup>1,2</sup>, Daniel Campo, PhD<sup>3</sup>, Tania B. Porras, MS<sup>1</sup>, Pushpinder Kaur, PhD<sup>1</sup>, Victoria A. Forte, MD<sup>4</sup>, Debu Tripathy, MD<sup>5</sup>, Janice Lu, MD, PhD<sup>6</sup>, Irene Kang, MD<sup>7</sup>, Michael F. Press, MD, PhD<sup>7</sup>, Young Ju Jeong, MD, PhD<sup>8</sup>, Anson Snow, MD<sup>1</sup>, Yue Zhu, BS<sup>1</sup>, Gabriel Zada, MD<sup>9</sup>, Naveed Wagle, MD<sup>6</sup>, and Julie E. Lang, MD<sup>1,10</sup>

<sup>1</sup>Division of Surgical Oncology, Department of Surgery and University of Southern California Norris Cancer Center, University of Southern California, Los Angeles, CA; <sup>2</sup>Department of Hematology and Medical Oncology, University Hospital Zurich, Zurich, Switzerland; <sup>3</sup>Department of Biological Sciences, University of Southern California, Los Angeles, CA; <sup>4</sup>Division of Medical Oncology, Department of Medicine, SUNY Downstate Medical Center, New York, NY; <sup>5</sup>Department of Breast Medical Oncology, UT MD Anderson Cancer Center, Houston, TX; <sup>6</sup>Division of Medical Oncology, Department of Medicine and University of Southern California Norris Cancer Center, University of Southern California, Los Angeles, CA; <sup>7</sup>Department of Pathology and University of Southern California Norris Cancer Center, University of Southern California, Los Angeles, CA; <sup>8</sup>Department of Surgery, Catholic University of Daegu School of Medicine, Daegu, Republic of Korea; <sup>9</sup>Department of Neurosurgery and University of Southern California Norris Cancer Center, University of Southern California, Los Angeles, CA; <sup>10</sup>Division of Breast Services, Department of General Surgery, Cleveland Clinic Breast Cancer Program, Cleveland, Ohio

## ABSTRACT

**Background.** Metastatic breast cancer (MBC) and the circulating tumor cells (CTCs) leading to macrometastases are inherently different than primary breast cancer. We evaluated whether whole transcriptome RNA-Seq of CTCs isolated via an epitope-independent approach may serve as a surrogate for biopsies of macrometastases.

**Methods.** We performed RNA-Seq on fresh metastatic tumor biopsies, CTCs, and peripheral blood (PB) from 19 newly diagnosed MBC patients. CTCs were harvested using the ANGLE Parsortix microfluidics system to isolate cells based on size and deformability, independent of a priori knowledge of cell surface marker expression.

**Results.** Gene expression separated CTCs, metastatic biopsies, and PB into distinct groups despite heterogeneity between patients and sample types. CTCs showed higher

expression of immune oncology targets compared with corresponding metastases and PB. Predictive biomarker ( $n = 64$ ) expression was highly concordant for CTCs and metastases. Repeat observation data post-treatment demonstrated changes in the activation of different biological pathways. Somatic single nucleotide variant analysis showed increasing mutational complexity over time.

**Conclusion.** We demonstrate that RNA-Seq of CTCs could serve as a surrogate biomarker for breast cancer macrometastasis and yield clinically relevant insights into disease biology and clinically actionable targets.

Metastatic breast cancer (MBC) is responsible for virtually all BC deaths. MBCs are often discordant in biomarker profiles when compared with the primary tumor.<sup>1</sup> The American Society of Clinical Oncology guidelines call for biopsies of metastases for biomarker testing to guide decision making for systemic therapy.<sup>2,3</sup> However, not all metastatic sites are amenable to safe percutaneous biopsy. Improved survival in MBC is in large

© The Author(s) 2022, corrected publication 2022

First Received: 3 September 2021

Accepted: 11 October 2021;

Published Online: 9 January 2022

J. E. Lang, MD

e-mail: LANGJ2@ccf.org

part due to the availability of targeted therapies.<sup>4,5</sup> Most patients with MBC unfortunately develop treatment resistance.<sup>6</sup>

Although molecular profiling of tumors may predict targeted therapy opportunities, primary tumor or single metastatic biopsy site driven approaches may not represent multiple non-overlapping oncogenic alterations driving biology in patients with multiple metastatic sites.<sup>7</sup>

Circulating tumor cells (CTCs) hold significant potential as liquid biopsies obtained via minimally invasive blood draws for the real-time assessment of a patient's tumor biology and heterogeneity.<sup>8,9</sup> CTCs have been shown to be prognostic in MBC and are present in 52–71% of MBC patients,<sup>10</sup> but have not provided predictive insights for targeted therapy. Hence, CTCs have not been used extensively to guide therapy decisions. A potential issue is the selection of cell populations based on cell surface marker expression, such as the only FDA approved method via the cell search system.<sup>11</sup> Another issue is that mere enumeration of CTCs or technological limitations have hampered the capability to interrogate CTC biology and gain insights into potentially targetable lesions.<sup>8</sup>

Many sequencing approaches available now, including those with clinical application, focus on DNA sequencing, but there are major concerns that not all DNA mutations are expressed.<sup>12</sup> Refinements in RNA-Seq technology now enable detailed molecular profiling of CTCs<sup>13</sup> beyond gene expression, offering the potential for predicting treatment options via a liquid biopsy.

We hypothesized that molecular characterization via whole transcriptome RNA sequencing of CTCs isolated in an unbiased, marker independent fashion can capture disease heterogeneity of MBC and may serve as a surrogate for the analysis of macrometastases to identify predictive biomarkers, potentially leading to new target discovery and explaining treatment resistance.

## MATERIALS AND METHODS

### *Study Design and Patient Population*

The project was designed as an observational study to evaluate whether RNA-sequencing (RNA-Seq) of CTCs can identify potential treatment targets. A total of 21 treatment naïve female MBC patients were prospectively enrolled at the Keck Medical Center and Norris Comprehensive Cancer Center at the University of Southern California (USC). Each patient underwent biopsies of macrometastases for clinical diagnostic purposes collected at baseline (prior to therapy for MBC) or upon disease progression prior to switching therapy. A baseline PB draw of 7.5 ml in an EDTA tube for CTC RNA-Seq was required

for inclusion. Data from 19 patient samples passing quality criteria were included in further analysis. Four of the 19 patients with progressive disease returned for repeat PB draws after approximately 6 months of treatment to track the changes in CTC biology over time. Response to therapy was assessed based on RECIST criteria.<sup>14</sup> All procedures, including written patient informed consent, were approved by the Institutional Review Board (IRB HS-14-00595 and HS-11-00208) at USC. This study was compliant with the REMARK criteria.<sup>15</sup>

### MOLECULAR MARKER-INDEPENDENT CTC ISOLATION

The Parsortix microfluidics filtration system (ANGLE plc, Surrey, United Kingdom) efficiently captures and highly enriches CTCs in a cell surface marker independent manner based on size and deformability,<sup>16–19</sup> reducing the number of contaminating white blood cells (WBCs) by roughly 5 orders of magnitude. The device has a Diagnostic Devices Directive CE Mark for clinical use in Europe. We have previously validated the capture efficiency of the device in our lab using breast cancer cells spiked into peripheral blood samples (Supplementary Fig. S1). A capture cassette with a critical gap of 10 microns was used to enrich CTCs.<sup>19</sup> Cell pellets were resuspended in 10  $\mu$ l of lysis buffer (NuGEN Technologies, Inc., San Carlos, CA) and stored at  $-80^{\circ}\text{C}$  for further use. Rigorous device cleaning was performed between samples. This cell surface marker independent approach allowed for the capture of heterogeneous CTC populations, including EpCAM negative cells and clusters of CTCs.<sup>17</sup> As processing time is critical to maximize capture efficiency<sup>20</sup> total time from blood draw to CTC harvest did not exceed 2 h. As negative controls, phosphate buffered saline (PBS) samples and PB from 5 healthy female donors were processed.

### *Sample Preparation and Whole Transcriptome RNA-Seq and Sanger Sequencing*

Either 50 ng of RNA from a metastasis or PB, isolated with a TRIzol or RiboPure kit (both Thermo Fischer Scientific, Waltham, MA), respectively, or 2  $\mu$ l of CTC lysate, were used to create cDNA for sequencing library preparation using the Ovation RNA-Seq System V2 and Ovation Ultralow Library System V2 (NuGEN Technologies, San Francisco, CA). Details regarding isolation and preparation of RNA can be found in Supplementary File S1. Sequencing was done on an Illumina HiSeq 2500 (Illumina, San Diego, CA) performing 100 base pair paired-end RNA-Seq using five samples per lane. Sanger sequencing was performed by Genewiz (South Plainfield, NJ, USA)

and the sequencing data was analyzed manually with 4Peaks (Nucleobytes, Aalsmeer, Netherlands) (Supplementary File S1, Supplementary Table S1). RNA-Seq data quality control and mapping were performed as previously described<sup>21</sup> (Supplementary File S1). For somatic SNV (single nucleotide variant) calling, the FASTQ files were processed following the Best Practices Workflow for variant calling with RNA-Seq from the Broad Institute (Supplementary File S1). The COSMIC database and 184 known driver genes in BC from the Integrative Onco Genomics database (<http://www.intogen.org/mutations/>)<sup>22</sup> were investigated for known SNVs in our data set. The driver gene analysis was done using Maftools.<sup>23</sup> We curated a list of 64 BC related genes with clinical and preclinical therapeutic, prognostic, or diagnostic implications, performing an extensive literature search<sup>24</sup> (Supplementary Table S2) representing breast cancer relevant pathways (EGFR/RAF/MEK, IGF-1/PI3K/AKT/mTOR, WNT/NOTCH/Hedgehog/FGF/MET, DNA damage repair, cell cycle, hormone receptor signaling, tumor suppressors, and tumor immunology). The FASTQ files, as well as the corresponding read count files for each sample were deposited in the Gene Expression Omnibus database (GSE113890).

### Statistical Analysis

Statistical analyses were conducted using GraphPad Prism (San Diego, CA, USA). For differences in gene expression, two-way ANOVA was used. For SNV comparison, the Wilcoxon sign-rank test and Friedman test were used. The number of uniquely mapped reads was compared using Kruskal-Wallis and Dunn's multiple comparison tests.

## RESULTS

### Patient Characteristic and Grouped Gene Expression Analysis Separates Sample Type

Table 1 lists all 19 patients with clinical annotations, including site of metastasis, biomarker [ER, progesterone receptor (PR), and HER2] expression as well as treatments received (two patient samples were excluded due to low read counts). The number of uniquely mapped reads was comparable between all sample groups (median uniquely mapped reads: CTCs  $30,962,730 \pm 19,535,767$ ; metastases  $40,580,689 \pm 14,659,648$ ; PB  $43,727,727 \pm 22,600,318$ ; CTCs vs PB  $p = 0.094$ ; CTCs vs metastases  $p = 0.094$ ; PB vs metastases  $p > 0.99$ . Median coverage: CTCs  $104X \pm 64$ ; metastasis  $139X \pm 64$ ; PB  $153X \pm 93$ ; CTCs vs PB  $p = 0.058$ ; CTCs vs metastasis  $p = 0.25$ ; PB vs metastasis

$p = 0.78$ ) (Supplementary Table S3). We obtained an average coverage higher than 50X for all but five of the samples, and coverage greater than 100X in 86% of the samples. The negative controls (PBS samples processed by Parsortix) yielded virtually no read counts (Supplementary Table S3).

Principal component analysis (PCA) showed separation of the majority of CTCs versus metastases and PB in PC1, and separation of CTCs and metastases from PB in PC2 (Fig. 1A). A Venn diagram is shown in Fig. 1B for inter-group comparison of gene expression in CTCs, metastases, and PB. A pairwise grouped comparison was used to identify the five most up- and downregulated genes with an expression change of at least 2-fold and adjusted  $p$ -value of  $< 0.05$ : CTCs vs PB downregulated: YTHDC1, CREG1, CLK2, ADIPOR1, RN7SL2; upregulated: GPRC5D, LINC01376, LOC727993, TAS1R3, ARC; CTCs + metastases vs PB downregulated: SNAP23, GALNS, SELENOS, RN7SL2, MORN1; upregulated MGP, GDF9, MYH11, AZGP1, LUM; CTCs vs metastases downregulated: ABCF1, TJP1, DLG5, PTPRK, H19; upregulated: GPRC5D, TMEM198, LOC727993, ARC, RNU6ATAC (Supplementary Fig. S2). In summary, these results show that RNA-Seq can detect distinct gene expression features in enriched CTCs compared with metastases and PB.

### Gene Expression of Potentially Clinically Actionable Genes Relevant to Breast Cancer

CTCs showed overall many more differentially expressed immune oncology target genes (Oncomine Immune Response Assay) compared with peripheral blood than did metastatic biopsies (overexpression: CTCs 131 vs metastases 15, 8.7-fold difference; downregulation: CTCs 38 vs metastasis 37, 1.03-fold difference). A total of 12 overexpressed and 15 downregulated genes compared with PB were in common between CTCs and metastasis (Fig. 2A, Supplementary Table S4). Notably, PD-L1 expression was significantly lower in both CTCs and metastases compared with PB (CTCs versus PB  $p = 3.5 \times 10^{-5}$ , CTCs vs metastases  $p = 0.004$  and metastases versus PB  $p = 0.004$ ) (Supplementary Fig. S3).

We found concordant expression of 50/64 (78%) potentially clinically actionable target genes in CTCs and corresponding metastases (Fig. 3A). No genes were uniformly overexpressed or downregulated in all sample groups (i.e., CTCs or metastases). Only 3/64 (4.7%) genes showed statistically significantly discordant expression in CTCs vs metastases (AKT3  $p = 0.018$ , CCND1  $p = 0.025$ , FOXA1  $p = 0.034$ ) (Fig. 3A). Figure 3B shows representative patient samples ( $n = 3$ ) for the expression of all 64 clinically actionable target genes with related clinical trials as well as targeted therapeutics. The majority of CTC and

TABLE 1 Clinical annotations of all patients

Patient	Primary tumor				Metastatic site				Overall treatment history	Therapy at time of metastasis diagnosis	CTC collection	
	Histology	ER	PR	HER2	Type	ER	PR	HER2			Therapy at visit 1	Therapy at visit 2
1	N/A	Y	N	N	N	Y	N	N	Everolimus/exemestane, paclitaxel/capecitabine/zoledronic acid, doxorubicin, gemcitabine	Doxorubicin	Doxorubicin	Gemcitabine
2	N/A	N/A	N/A	N/A	LN	Y	Y	N	ALND + XRT, tamoxifen, anastrozole/lanreotide	None	None	
3	IDC	N	N	N	Brain	N	N	N	Segmental mastectomy/ SLN+ IORT, docetaxel/ cyclophosphamide, XRT, capecitabine/ denosumab	Capecitabine	Capecitabine	
4	PLC	Y	Y	N	Ascites	N/A	N/A	N/A	Anastrozole/palbociclib, paclitaxel, doxorubicin	Doxorubicin	Doxorubicin	
5	N/A	Y	Y	N	Liver	Y	Y	N	Chemotherapy (unknown)/ XRT, tamoxifen, paclitaxel/cisplatin, anastrozole, palbociclib/ letrozole	Letrozole + palbociclib	Letrozole + palbociclib	
6	N/A	Y	N	N	Brain	Y	N	N	Chemotherapy (unknown) + right mastectomy, anastrozole	Anastrozole	Anastrozole	
7	IDC	Y	Y	N	Breast	Y	Y	N	Lumpectomy	None	None	
8	N/A	Y	Y	N	Ascites	N/A	N/A	N/A	Lumpectomy/ chemotherapy (unknown)/ radiation, carboplatin/paclitaxel, anastrozole, fulvestrant/ palbociclib, capecitabine, doxorubicin, everolimus/ exemestane	Doxorubicin	Doxorubicin + gemcitabine	Gemcitabine + exemestane + everolimus
9	IDC	Y	Y	N	Pleural effusion	Y	N	N	Dose dense adriamycin/ cyclophosphamide, MRM, paclitaxel, tamoxifen/ XRT, anastrozole, leuprolide/ zoledronic acid, fulvestrant/ leuprolide depot, capecitabine, everolimus/ exemestane, paclitaxel, doxorubicin	Paclitaxel > doxorubicin	Paclitaxel > doxorubicin	
10	IDC	Y	Y	N	Breast	Y	Y	N	Doxorubicin/ cyclophosphamide/ paclitaxel, left segmental mastectomy with lymphadenectomy, tamoxifen/ leuprolide/ zoledronic acid + XRT, exemestane/ leuprolide/ zoledronic acid, fulvestrant/ palbociclib, capecitabine + XRT,	Capecitabine	Capecitabine + XRT	Paclitaxel
11	IDC	N	N	N	CSF	N	N	N	Adriamycin/cyclophosphamide, paclitaxel, right MRM + ALND + XRT, right craniotomy w/ removal right occipital mass + cytarabine, carboplatin/ capecitabine	Carboplatin + capecitabine	Carboplatin + capecitabine	
12	ILC	N	N	N	Bone	N/A	N/A	N/A	Segmental mastectomy, brachytherapy, docetaxel/ cyclophosphamide, doxorubicin/cyclophosphamide/ paclitaxel + XRT, anastrozole, letrozole, capecitabine	Capecitabine, letrozole	Capecitabine + letrozole	
13	IDC+IBC	N	N	N	Brain	N	N	Y	Simple mastectomy, carboplatin/taxotere + radical mastectomy,	None	None	

Table 1 (continued)

Patient	Primary tumor		Metastatic site				Overall treatment history	Therapy at time of metastasis diagnosis	CTC collection			
	Histology	Histology	ER	PR	HER2	Type			ER	PR	HER2	Therapy at visit 1
14	IDC	IDC	Y	Y	N	Breast	Y	N	N	gencitabine/trastuzumab, surgical excision by suboccipital craniotomy + gamma knife Adriamycin/ cyclophosphamide, paclitaxel, tamoxifen, mastectomy + XRT, leuprolide/ anastrozole	Leuprolide + anastrozole	Leuprolide + anastrozole
15	IDC	IDC	Y	Y	N	Pleural effusion	Y	N	N	Adriamycin/ cyclophosphamide /docetaxel, BMX + bilateral ALND, capecitabine + XRT, tamoxifen, leuprolide/ zoledronic acid, capecitabine, gencitabine/ carboplatin	Capecitabine	Capecitabine
16	IDC	IDC	N/A	N/A	N/A	Pleural effusion	N	N	N	Adriamycin/ cyclophosphamide, paclitaxel, tamoxifen, mastectomy	None	None
17	N/A	N/A	Y	Y	N	CSF	Y	N	N	Mastectomy/ adjuvant chemotherapy/ XRT + adjuvant hormonal therapy, anastrozole, fulvestrant, paclitaxel/ capecitabine, carboplatin, vinorelbine	Gencitabine	Vinorelbine
18	IDC+ILC	IDC	Y	Y	N	Skin	Y	N	N	Letrozole + XRT	XRT + letrozole	Letrozole
19	IDC	IDC	Y	Y	Y	Pleural effusion	N/A	N/A	N/A	Docetaxel/ trastuzumab/ pertuzumab, trastuzumab/ pertuzumab/ letrozole	Docetaxel + Trastuzumab + pertuzumab	Docetaxel + Trastuzumab + pertuzumab

ALND: axillary node dissection; CSF: cerebrospinal fluid; IBC: inflammatory breast cancer; IDC: invasive ductal carcinoma; ILC: invasive lobular carcinoma; IORT: intraoperative radiation therapy; PLC: pleomorphic lobular carcinoma; LN: lymph node; MRM: modified radical mastectomy; SLN: sentinel lymph nodes; WBRT: whole brain radiation therapy; XRT: radiotherapy; BMX: bilateral mastectomy; N/A: information not available.

metastasis samples showed overexpression of these targetable genes compared with PB, with few exceptions (i.e., lower or weak expression, < 2-fold) (Fig. 3B) (Supplementary Fig. S5, Supplementary Table S1). These results indicated that CTCs could potentially serve as a surrogate for distant macrometastases for the identification of drug-gable targets.

*Sequential Analysis of CTC Samples*

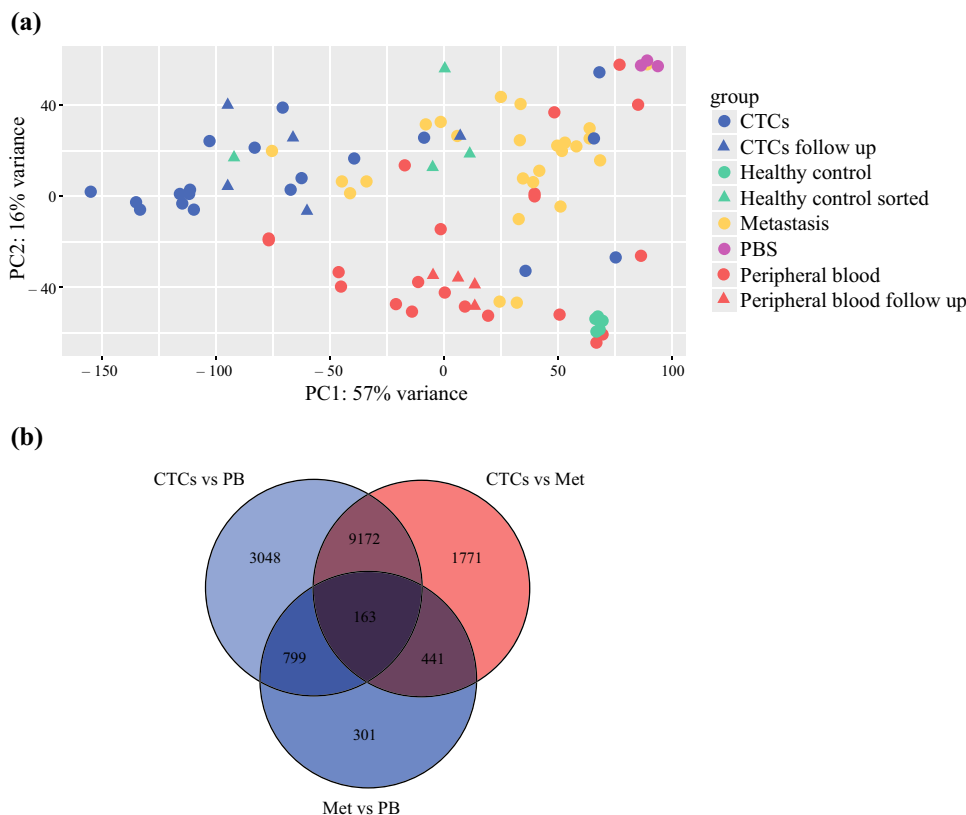
We tracked four patients with repeated harvest of CTCs at a second time point, as well as obtaining the imaging studies and systemic therapies these patients received (average time between first and second CTC harvest time point was  $4 \pm 0.8$  months). Representative results for two patients are shown in Fig. 4 (the remaining data can be found in Supplementary Fig. S6). The metastatic sites profiled for these patients were pleural effusions (ER/PR+, HER2-) (patient 1 in Table 1) (Fig. 4A and B), (ER+, PR/HER2-) (patient 15 in Table 1) (Fig. 4C and D). For the first patient, hormone receptor genes were downregulated upon sequential CTC assessment after receiving tamoxifen (Fig. 4B) (fold change expression compared with PB: receptors—metastasis  $1.29 \pm 0.94$ , CTCs  $2.56 \pm 1.96$ , CTCs follow-up  $-0.2 \pm 1.48$ ; cell cycle—metastasis  $-0.71 \pm 2.59$ , CTCs  $0.83 \pm 1.7$ , CTCs follow-up  $-2.50 \pm$

$2.59$ ; EGFR signaling—metastasis  $-0.73 \pm 2.49$ , CTCs  $-0.43 \pm 1.71$ , CTCs follow-up  $-2.39 \pm 3.31$ ). The difference in expression was statistically significant for comparison of CTCs and CTCs follow-up for receptor expression ( $p = 0.014$ ) and cell cycle gene expression ( $p = 0.0005$ ). The follow-up CTC sample in the second patient showed upregulation of DNA-damage repair genes under treatment with an alkylating agent (doxorubicin) (Fig. 4D) (fold change expression compared to PB: receptors—metastasis  $3.90 \pm 1.48$ , CTCs  $3.54 \pm 2.98$ , CTCs follow-up  $3.68 \pm 1.38$ ; cell cycle—metastasis  $2.24 \pm 2.47$ , CTCs  $3.27 \pm 2.55$ , CTCs follow-up  $-0.59 \pm 3.97$ ; DNA damage repair—metastasis  $-0.33 \pm 0.92$ , CTCs  $0.39 \pm 1.42$ , CTCs follow-up  $0.84 \pm 1.99$ ). For both patients, the CTC follow-up samples showed a reduction in cell cycle gene expression. These data revealed an evolution of biological features within each patient’s disease under therapeutic pressure, with implications for clinical management.

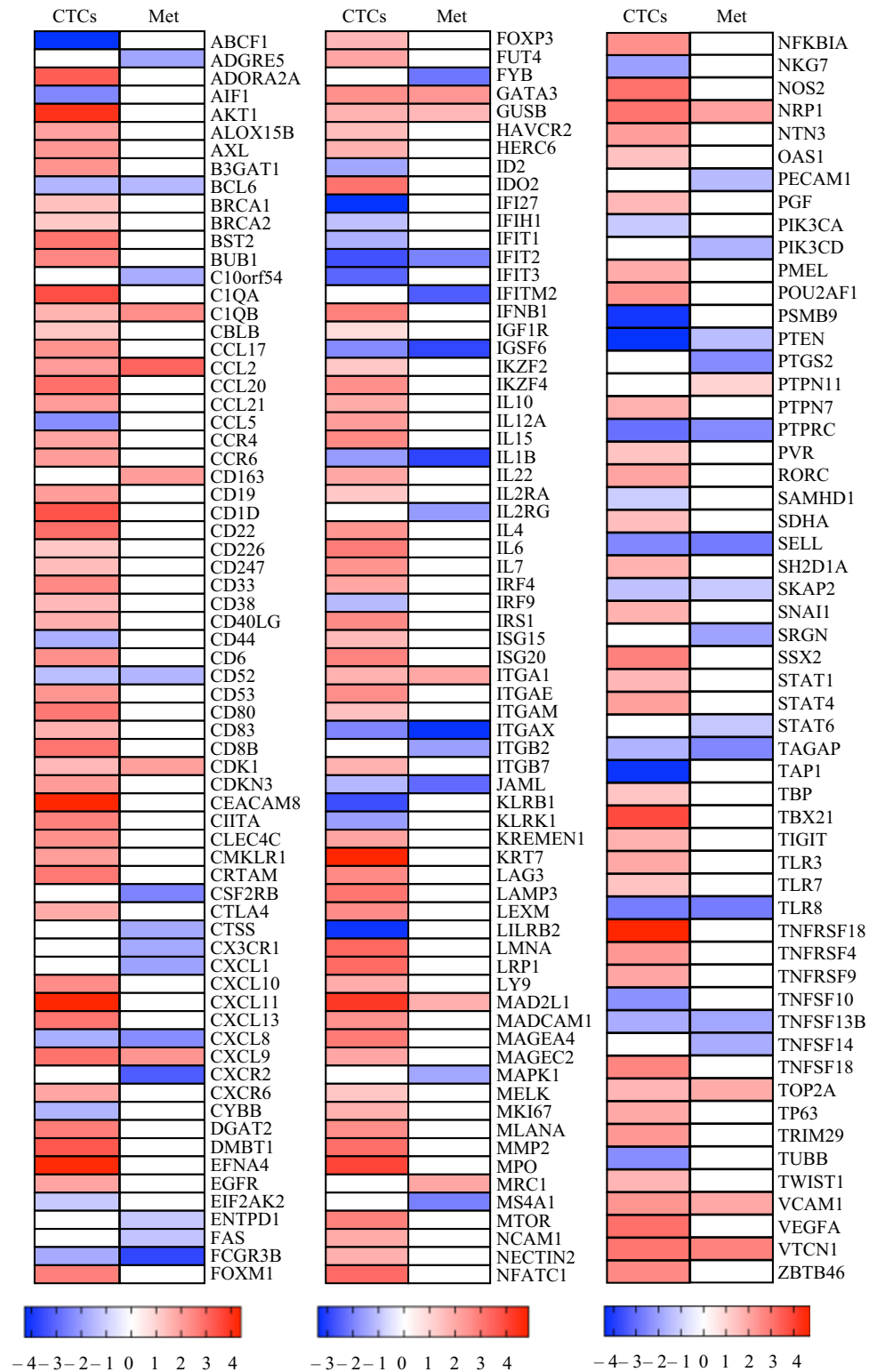
*Somatic Single Nucleotide Variant (SNV) Analysis*

Across all samples, we detected SNVs in 1754 genes (1608 in CTCs, 212 in metastases). At the gene level, 65/212 (31%) of the gene mutations in the metastatic biopsies were also present in the CTCs (Fig. 5A, Supplementary Table S5). A total of 2258 somatic mutations were

**FIG. 1** Principal component (PC) analysis and differential gene expression distinguishes different sample types (CTCs, CTCs follow-up, metastases, PB, PB follow-up, healthy controls, healthy controls sorted, PBS). **A** Principal component analysis demonstrating separation of the majority of CTCs versus metastases and PB in PC1 (57% of variance explained), and separation of CTCs and metastases from PB in PC2 (16% of variance explained). **B** Venn diagram for inter-group comparison of differentially expressed genes (FDR adjusted  $p < 0.05$ )



**FIG. 2** Absolute normalized expression (RPKM) of immune response assay genes in CTCs, metastases and PB. Heatmap showing the expression of immune target genes ( $n = 207$ ) (log fold change) from the Oncomine Immune Response Research Assay in CTCs and metastases normalized to PB based on RNA-Seq data



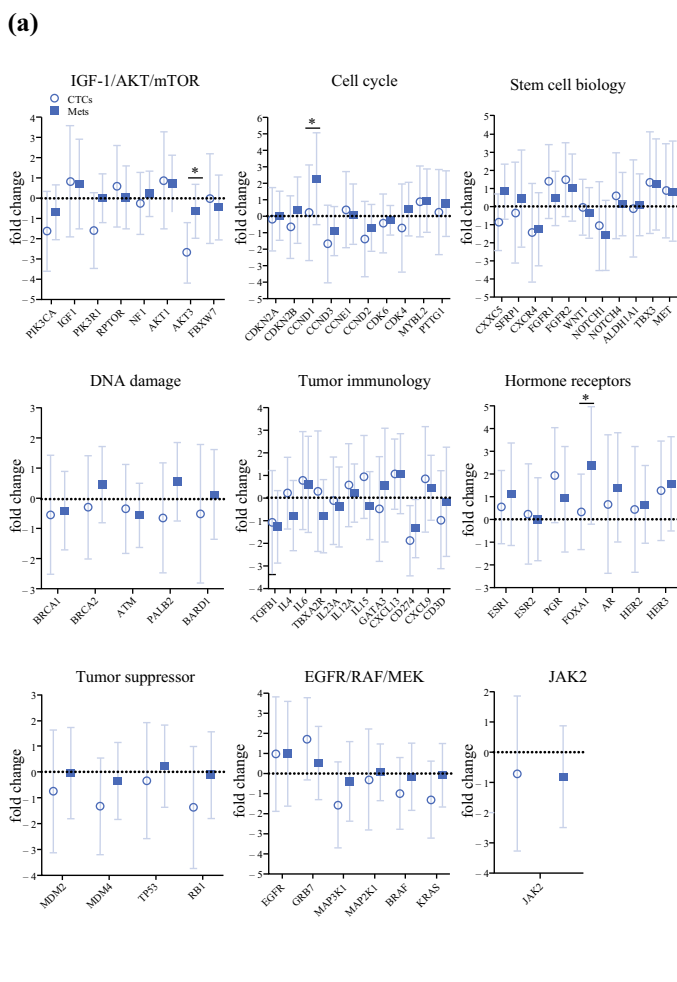
found across all samples, with CTCs showing a 9.4-fold higher number of SNVs compared with metastases (2041 in CTCs, 217 in metastases; mean and SD: CTCs  $93 \pm 231$  vs metastases  $13 \pm 21$ ,  $p = 0.01$ ) (Fig. 5B). We detected 344 variants (17%) found in all CTCs and 42 variants (1.9%) in

all metastases that corresponded to SNVs found in the COSMIC (Catalogue of Somatic Mutations in Cancer) database (Supplementary Table S6). Our data showed increasing genomic complexity represented by a higher number of SNVs in 3/4 patients with follow-up at the

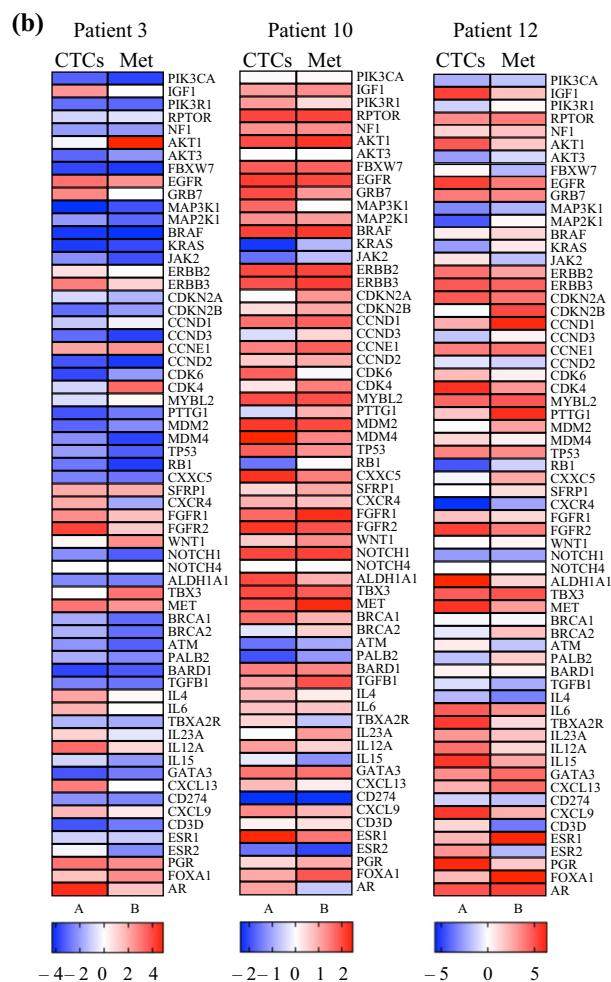
second time point (mean SNVs: metastasis  $4 \pm 1$ , CTCs  $35 \pm 20$ , CTCs follow-up  $110 \pm 119$ , CTCs vs CTCs follow-up  $p = 0.035$ ) (Fig. 5C and D, Supplementary Fig. S7A–C). Sanger sequencing validated 6 of 10 selected RNA-Seq variants (60%) (Supplementary Fig. S7). We ranked the top 20 most frequently mutated genes across our samples: AHNAK, ALMS1, ANKRD12, ARID1A, ARHGAP35, BEST1, BPTF, CALM2, F5, HIVEP1, MACF1, MDN1, MIK67, MUC3A, MUC12, MUC16, SOS1, TET2, WIPF1, ZFH4 (Fig. 5E, Supplementary Fig. S9, Supplementary Table S8) and validated these genes in publicly available data sets using cBioPortal, comparing metastatic ( $n = 396$ ) vs non-metastatic ( $n = 5158$ ) BC samples. Seventeen out of those 20 genes were mutated in BC with VAF varying from 0.2 to 9% (Fig. 5F). We found a significant difference in SNVs between metastatic and non-metastatic cases ( $p = 0.009$ ) (mean metastatic vs non-metastatic  $1.94 \pm 1.84$  vs

**FIG. 4** Intra-patient analysis predictive biomarker gene expression in metastatic biopsy tissue and two serial CTC harvest time points. **A**, **C** Clinical data (including treatment and CT imaging studies). **B**, **D** Representative clinically actionable genes and pathways (3 per patient) for CTCs: CTC follow-up samples for 2 patients are shown. The gray bar indicates the mean expression and standard deviation for all genes represented in the biological function. Expression of individual genes are represented by the black squares. A two-way ANOVA was performed to compare overall expression between all three sample types (metastatic sample, CTCs, CTCs follow-up) ( $*p = 0.014$ ,  $***p = 0.0005$ )

$1.2 \pm 2.35$ ), with up to 10-fold differences for single genes (i.e., TET2). Comparing 184 putative BC driver genes identified in the IntOGen-mutations platform<sup>22</sup> we found SNVs in 44/184 (24%) in our RNA-Seq data set (Supplementary Table S9). Of these, 34/44 (77.3%) were present

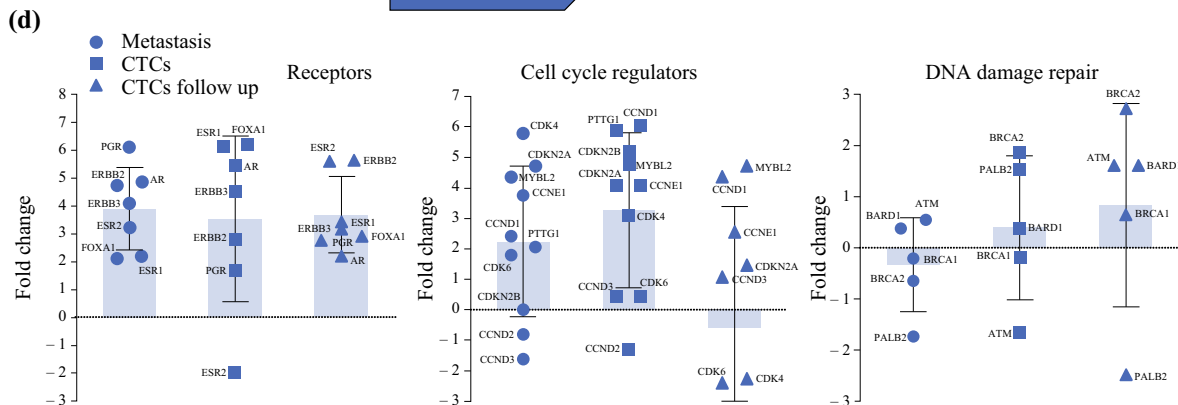
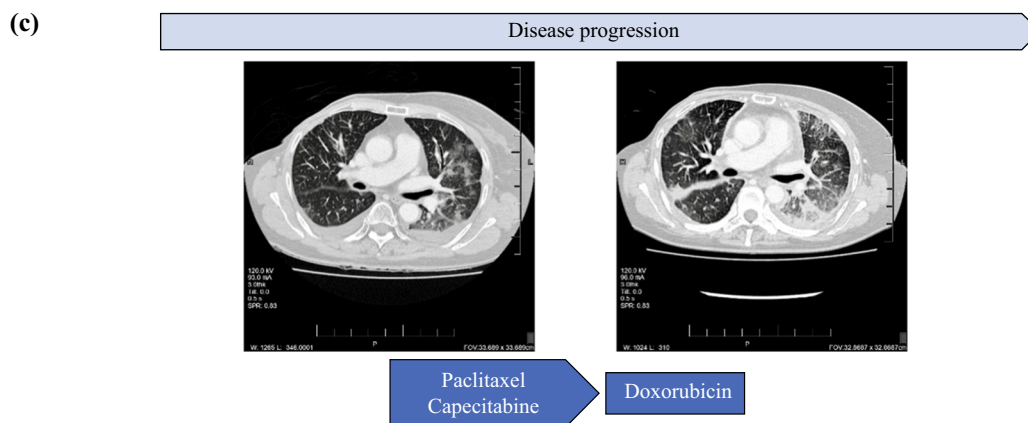
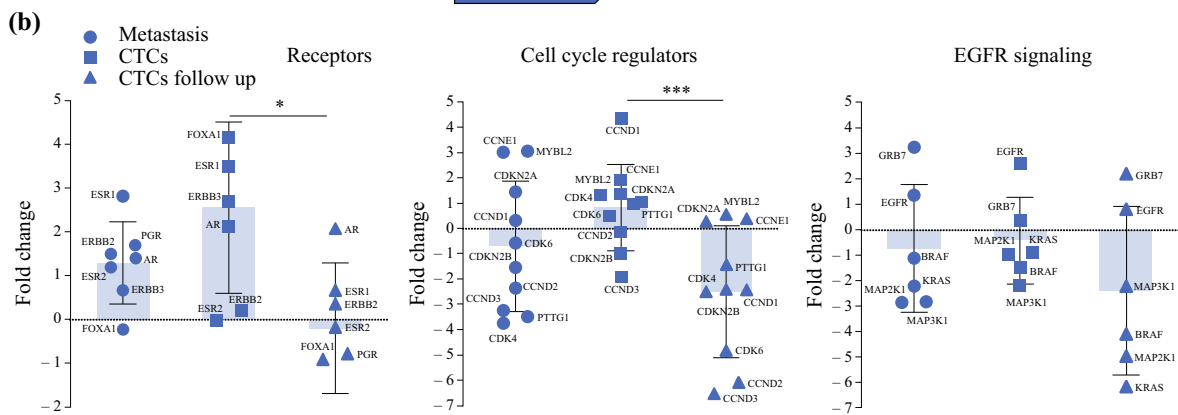
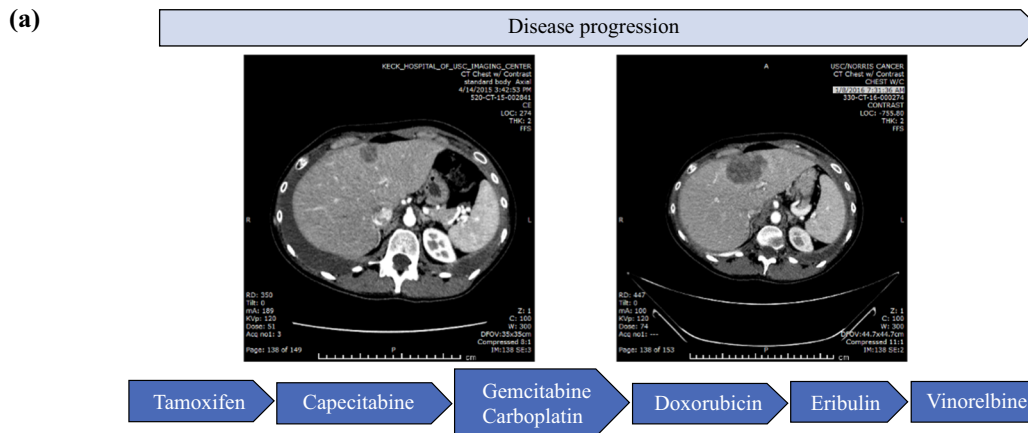


**FIG. 3** Comparison of potentially clinically actionable breast cancer related genes in CTCs and metastases normalized to PB. **A** Comparison of gene expression between paired CTCs and metastases in 64 clinically actionable genes grouped into 9 biological pathways for 19 patients ( $*p \leq 0.05$ ) (2-way ANOVA).



**B** Expression of 64 clinically actionable target genes in three representative patient samples (patient numbers correspond to Table 1): expression is shown as log fold change values for each gene in CTCs and matched metastatic biopsies, each normalized to the patient's own PB





only in CTC samples, 4/44 (9.1%) only in metastatic samples and 5/44 (11.4%) in both CTCs and metastases.

In summary, our analysis demonstrated a higher number of heterogeneous somatic mutations in CTCs compared with macrometastatic biopsies, an increase in the number of SNVs in CTCs over time, and revealed that RNA-Seq of CTCs can detect driver gene mutations in MBC.

## DISCUSSION

We present the gene expression profiling of enriched CTCs from MBC patients with comparison to metastases and PB, all acquired prior to treatment or at disease progression prior to a new line of therapy. PCA analysis showed that all sample groups (CTCs, metastases, and PB) separated in PC1, with most CTCs and metastases clustering together in PC2. The partial overlap with PB might be explained by findings that CTCs frequently associate with WBCs, in particular neutrophils.<sup>25</sup> Our lab has also previously shown that even ultra-pure CTC populations express WBC genes.<sup>21</sup> Both metastatic and CTC samples (expected 4–10 background leukocytes per CTC after Parsortix enrichment) likely also contain WBCs.

Several genes with biological and clinical implications for BC were highly expressed in CTCs. GPRC5D has been previously associated with tamoxifen resistance.<sup>26</sup> TMEM198 promotes LRP6 phosphorylation in activating Wnt signaling,<sup>27</sup> which has been associated with CSC biology in BC. The apoptosis inhibitor ARC has been associated with chemotherapy resistance, tumorigenesis, and metastasis in the polyoma middle T-antigen (PyMT) transgenic mouse model of BC.<sup>28</sup> It has also been shown to lead to TP53 inactivation in TP53 WT malignancies.<sup>29</sup> LOC727993, a non-coding antisense RNA of the gene known as PDYN-AS1, and RNU6ATAC, a small nuclear RNA associated with U12-dependent splicing, have not previously been demonstrated to be involved with tumor biology. Our approach identified both known and novel genes associated with CTCs, suggesting that CTCs might be suitable as a discovery tool to better understand the fundamental tumor biology of metastasis.

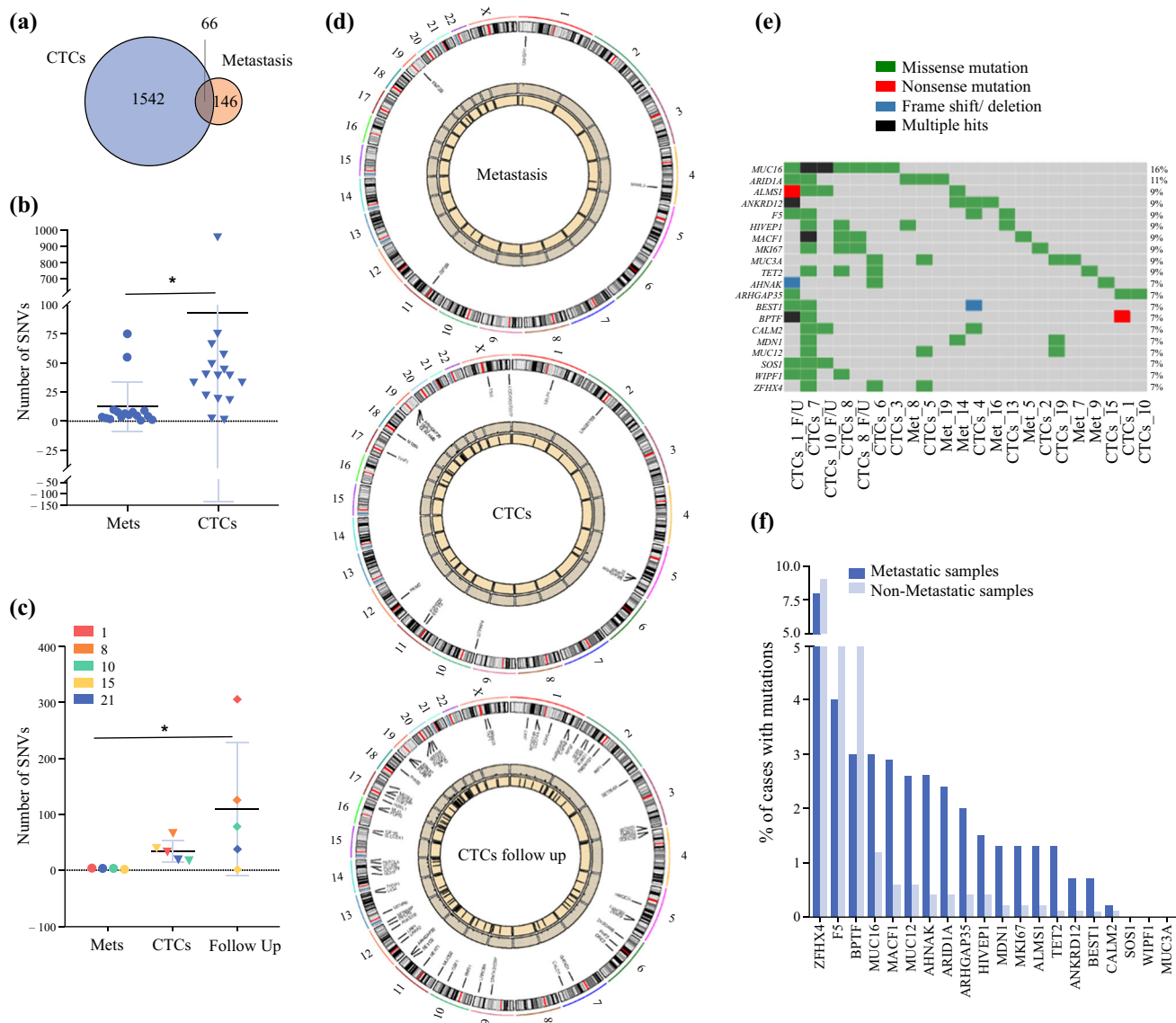
We found highly concordant expression in potentially clinically actionable genes in corresponding CTC and metastatic samples, demonstrating the potential clinical relevance of CTCs as predictive biomarkers in BC. Nevertheless, we also observed discordant results, which could be due to various conditions: (1) a heterogeneous origin of CTCs from various metastatic sites or seeding of CTCs from the primary tumor site, (2) changes in transcriptional programs once cells “settle” in a new environment, influenced by tissue or site-specific micro-environmental cues,<sup>30</sup> or (3) differences in the timing of when the

metastatic (more remote and established) vs seeding of CTCs occurs (which may be more reflective of recent genomic alterations and treatments).

Longitudinal analysis of four patients with serial CTC assessments showed changes in biological pathway activation during treatment and disease progression. We found markedly increased genetic complexity in 3 out of 4 patients over time. These results indicate that serial CTC harvest might capture changes in additional mutation burden as a cancer evolves, particularly under the selection pressure of anti-cancer therapies.<sup>31</sup> Periodic surveying of the mutational evolution using CTCs could thus impact clinical decision making. Additionally, CTCs might capture the mutational landscape of a patient’s cancer from different metastatic sites more comprehensively than single site biopsies.<sup>32</sup>

Compared with gene expression of potentially actionable genes, we observed a much lower concordance in our SNV analysis. As there is no standard tool or pipeline for using RNA-Seq to call SNVs, we established a workflow for the purpose. SNV calling decreases with low read depth or low allelic frequency, diminishing the sensitivity of SNV detection.<sup>33</sup> The lack of overlap and greater genomic complexity in CTCs could also represent the pool of heterogeneous somatic mutations from various metastatic sites and different cancer cell clones compared with individual metastatic sites. The strength of our approach is the inference of expressed mutations, given that not all DNA mutations are expressed. The finding that lncRNAs were frequently mutated in CTCs and metastases offers an interesting opportunity to further investigate the role of regulatory RNAs in metastasis.<sup>34</sup> As this aspect of our paper is the most speculative piece of the manuscript, we believe that better tools are needed for more sensitive SNV calling from complex samples such as circulating tumor cells. Single cell sequencing studies may shed light on this by controlling for the input of cancer vs peripheral blood mononuclear cells.

Analysis of the expression of IO target genes demonstrated that CTCs expressed 170/200 genes related to immune response while metastatic biopsies expressed only 52/200 such genes, suggesting an important role for the immune system in CTC biology and potential immune escape. During dissemination, CTCs are exposed to many types of stress in the blood microenvironments and direct exposure to immune surveillance. Our results are in line with previous studies demonstrating upregulation of potential immune-escape mechanisms.<sup>35</sup> Several highly expressed IO genes in CTCs in our study have been shown to play important roles in immune evasion and metastatic efficiency: AKT1 can potentially suppress immunodetection by activating myeloid suppressor cells.<sup>36</sup> The complement component C1q might facilitate the metastatic



**FIG. 5** SNVs detected in CTCs and metastases compared with PB. **A** Overlap of all genes with SNVs found across all CTCs and metastases samples. **B** Number of SNVs in individual metastatic samples and CTCs side-by-side (\* $p = 0.011$ ) (Wilcoxon test). **C** Number of SNVs in 5 samples with CTC follow-up data (patient numbers correspond to Table 1). Each *column* represents a sample type, each *color* represents a patient with ID number (\* $p = 0.02$ ) (Friedman test) (the SNV for metastasis sample in patient 8 failed). **D** Circos plots showing mutations and corresponding gene expression for one patient (corresponding to patient 10 in Table 1) including CTCs, CTCs follow-up and metastasis. *Innermost circle*: position of

potential of CTCs.<sup>37</sup> CXCL9-11 might act as a double-edged sword via paracrine and autocrine signaling or interaction with PD-L1, inhibiting or facilitating immune escape and metastatic seeding, respectively.<sup>38</sup> We found expression of PD-L1 to be lower in CTCs and metastases compared with PB samples in our study. Although this marker has been suggested as a potential biomarker in

the mutated gene on the corresponding chromosome on the *outermost circle*. *Middle circle*: RPKM value for expression of genes with SNVs. **E** Oncoplot showing driver genes ( $n = 20$ ) with the highest number of mutations across the highest number of samples (percentage of samples with mutations on *right* side of the graph). **F** Validation of mutated driver genes identified by RNA-Seq in publicly available data sets using cBioPortal comparing metastatic ( $n = 396$ ) vs non-metastatic ( $n = 5158$ ) breast cancer data sets (mean number of mutations in metastatic vs non-metastatic samples was  $1.94 \pm 1.84$  vs  $1.20 \pm 2.35$ ;  $p = 0.009$ ) (Wilcoxon test)

CTCs,<sup>39</sup> our study differs regarding the detection method and CTC capture platform with consideration of gene expression relative to background PB. High gene expression of PD-L1 in PBMCs is to be expected; the *Human Protein Atlas* shows high expression of CD274 (PD-L1) in

basophils.<sup>40</sup> These findings have implications for the use of immune targeted drugs and warrant further investigation into immune targeting of CTCs.<sup>41</sup>

There are several limitations of our study such as a relatively small number of patients, and CTC enrichment purity. We applied per patient normalization, utilizing matched white blood cells as a pre-specified analysis plan regarding our primary research question of comparing the gene expression of CTCs vs metastases for a list of well characterized, potentially clinically actionable marker genes. Our strategy of per patient normalization to PB signal is novel, focusing attention on genes with strong differential expression between tumor and blood by controlling for leukocyte background.<sup>21,42</sup> Thus, subtle differences in gene expression might not be captured with our method. Ideally, sequencing of pure cell populations, even at single cell level should be attempted to characterize differences of gene expression between CTCs and WBCs more stringently.<sup>43</sup>

For gene expression results, standard normalization (reported as reads per kilobase of transcript, per million mapped (RPKM)) was applied but, due to the nature of our approach, we cannot rule out a certain degree of amplification bias. However, we previously utilized unspiked negative controls and extensively validated our RNA amplification strategy.<sup>21,44</sup> Although we successfully detected SNVs, our current data analysis pipeline does not allow for the detection of copy number variation. SNV-calling from RNA-Seq is less well established compared with the DNA based method, and further validation will be needed in the future.

## CONCLUSIONS

RNA-Seq of Parsortix-enriched CTCs could lead to minimally invasive, real-time diagnostic strategies for precision therapeutic decision making for MBC patients. Our approach could serve as a surrogate liquid biopsy for potentially clinically actionable drug target gene expression and mutations, allowing longitudinal assessment of the evolution of a patient's cancer.

**FUNDING** Funding was provided by ANGLE plc. The project described was supported in part by Award Number P30CA014089 from the National Cancer Institute. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Cancer Institute or the National Institutes of Health. ANGLE plc., Personal Grant to J.E.L., National Cancer Institute, P30CA014089.

**DISCLOSURES** The authors declare no competing interests for the presented study. Although ANGLE plc provided research funds to the author's institution, the authors have no relevant financial disclosures. The study sponsor did not influence data acquisition, analysis or the

decision to publish. Debasish Tripathy: Research support (paid to institution): Novartis, Pfizer, Polyphor Consulting fees (paid to self): AstraZeneca, GlaxoSmithKline, Gilead, Exact Sciences, OncoPep, Novartis, Pfizer. Irene Kang: Puma Biotechnology – speaker's bureau, consulting fees Bristol Myers Squibb — consulting fees. Michael Press: Consulting or advisory role: AstraZeneca, Biocartis, Cepheid, Inc.; Eli Lilly & Company, Merck & Co., Puma Biotechnology; Expert testimony: Amgen, Inc.; Private equity: TORL Biotherapeutics, LLC.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1245/s10434-021-11135-2>.

**OPEN ACCESS** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## REFERENCES

- Vignot S, et al. Discrepancies between primary tumor and metastasis: a literature review on clinically established biomarkers. *Crit Rev Oncol Hematol*. 2012;84(3):301–13.
- Hammond ME, et al. American Society of Clinical Oncology/College of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer (unabridged version). *Arch Pathol Lab Med*. 2010;134(7):e48–72.
- Wolff AC, et al. Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update. *J Clin Oncol*. 2013;31(31):3997–4013.
- Robson M, et al. Olaparib for metastatic breast cancer in patients with a germline BRCA mutation. *N Engl J Med*. 2017;377(6):523–33.
- Murthy RK, et al. Tucatinib, trastuzumab, and capecitabine for HER2-Positive metastatic breast cancer. *N Engl J Med*. 2020;382(7):597–609.
- Redig AJ, McAllister SS. Breast cancer as a systemic disease: a view of metastasis. *J Intern Med*. 2013;274(2):113–26.
- Wheler JJ, et al. Cancer therapy directed by comprehensive genomic profiling: A single center study. *Cancer Res*. 2016;76(13):3690–701.
- Castro-Giner F, Aceto N. Tracking cancer progression: from circulating tumor cells to metastasis. *Genome Med*. 2020;12(1):31.
- Diamantopoulou Z, Castro-Giner F, Aceto N. Circulating tumor cells: Ready for translation? *J Exp Med*, 2020. **217**(8)
- Cristofanilli M, et al. Circulating tumor cells, disease progression, and survival in metastatic breast cancer. *N Engl J Med*. 2004;351(8):781–91.
- Banko P, et al. Technologies for circulating tumor cell separation from whole blood. *J Hematol Oncol*. 2019;12(1):48.

12. Siegel MB, et al. Integrated RNA and DNA sequencing reveals early drivers of metastatic breast cancer. *J Clin Invest.* 2018;128(4):1371–83.
13. Islam S, et al. Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq. *Genome Res.* 2011;21(7):1160–7.
14. Eisenhauer EA, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer.* 2009;45(2):228–47.
15. McShane LM, et al. Reporting recommendations for tumour MARKer prognostic studies (REMARK). *Br J Cancer.* 2005;93(4):387–91.
16. Hvichia GE, et al. A novel microfluidic platform for size and deformability based separation and the subsequent molecular characterization of viable circulating tumor cells. *Int J Cancer.* 2016;138(12):2894–904.
17. Gkoutela S, et al. Circulating tumor cell clustering shapes DNA methylation to enable metastasis seeding. *Cell.* 2019;176(1–2):98–112.
18. Lampignano R, et al. a novel workflow to enrich and isolate patient-matched EpCAM(high) and EpCAM(low/negative) CTCs enables the comparative characterization of the PIK3CA status in metastatic breast cancer. *Int J Mol Sci.* 2017;18(9):1885.
19. Miller MC, et al. The Parsortix Cell Separation System-A versatile liquid biopsy platform. *Cytometry A.* 2018;93(12):1234–9.
20. Ring A, et al. EpCAM based capture detects and recovers circulating tumor cells from all subtypes of breast cancer except claudin-low. *Oncotarget.* 2015;6(42):44623–34.
21. Lang JE, et al. RNA-seq of circulating tumor cells in stage II-III breast cancer. *Ann Surg Oncol.* 2018;25(8):2261–70.
22. Gonzalez-Perez A, et al. IntOGen-mutations identifies cancer drivers across tumor types. *Nat Methods.* 2013;10(11):1081–2.
23. Mayakonda A, et al. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res.* 2018;28(11):1747–56.
24. Kaur P, et al. Comparison of TCGA and GENIE genomic datasets for the detection of clinically actionable alterations in breast cancer. *Sci Rep.* 2019;9(1):1482.
25. Szczerba BM, et al. Neutrophils escort circulating tumour cells to enable cell cycle progression. *Nature.* 2019;566(7745):553–7.
26. Mendes-Pereira AM, et al. Genome-wide functional screen identifies a compendium of genes affecting sensitivity to tamoxifen. *Proc Natl Acad Sci USA.* 2012;109(8):2730–5.
27. Liang J, et al. Transmembrane protein 198 promotes LRP6 phosphorylation and Wnt signaling activation. *Mol Cell Biol.* 2011;31(13):2577–90.
28. Medina-Ramirez CM, et al. Apoptosis inhibitor ARC promotes breast tumorigenesis, metastasis, and chemoresistance. *Cancer Res.* 2011;71(24):7705–15.
29. Foo RS, et al. Regulation of p53 tetramerization and nuclear export by ARC. *Proc Natl Acad Sci USA.* 2007;104(52):20826–31.
30. Gupta GP, et al. Identifying site-specific metastasis genes and functions. *Cold Spring Harb Symp Quant Biol.* 2005;70:149–58.
31. Venkatesan S, et al. Treatment-induced mutagenesis and selective pressures sculpt cancer evolution. *Cold Spring Harb Perspect Med.* 2017;7(8):a026617.
32. Yates LR, et al. Genomic evolution of breast cancer metastasis and relapse. *Cancer Cell.* 2017;32(2):169–184 e7.
33. Coudray A, et al. Detection and benchmarking of somatic mutations in cancer genomes using RNA-seq data. *PeerJ.* 2018;6:e5362.
34. Schmitt AM, Chang HY. Long noncoding RNAs in cancer pathways. *Cancer Cell.* 2016;29(4):452–63.
35. Steinert G, et al. Immune escape and survival mechanisms in circulating tumor cells of colorectal cancer. *Cancer Res.* 2014;74(6):1694–704.
36. Jia A, et al. The kinase AKT1 potentiates the suppressive functions of myeloid-derived suppressor cells in inflammation and cancer. *Cell Mol Immunol.* 2021;18(4):1074–6.
37. Bulla R, et al. C1q acts in the tumour microenvironment as a cancer-promoting factor independently of complement activation. *Nat Commun.* 2016;7:10346.
38. Tokunaga R, et al. CXCL9, CXCL10, CXCL11/CXCR3 axis for immune activation - A target for novel cancer therapy. *Cancer Treat Rev.* 2018;63:40–7.
39. Mazel M, et al. Frequent expression of PD-L1 on circulating breast cancer cells. *Mol Oncol.* 2015;9(9):1773–82.
40. Uhlen M, et al. Proteomics. Tissue-based map of the human proteome. *Science.* 2015;347(6220):1260419.
41. Christofi T et al. Current perspectives in cancer immunotherapy. *Cancers (Basel),* 2019. **11**(10)
42. Peixoto BR, et al. Evaluation of reference-based two-color methods for measurement of gene expression ratios using spotted cDNA microarrays. *BMC Genomics.* 2006;7:35.
43. Razavi P, et al. High-intensity sequencing reveals the sources of plasma circulating cell-free DNA variants. *Nat Med.* 2019;25(12):1928–37.
44. Porras TB, et al. Challenges in using liquid biopsies for gene expression profiling. *Oncotarget.* 2018;9(6):7036–53.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.