Single-Cell Resolution Characterization of Circulating Tumor Cell Clusters

Inauguraldissertation

zur
Erlangung der Würde eines Doktors der Philosophie
vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel

von

Barbara Maria Szczerba

aus Polen

2019

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät
auf Antrag von

Prof. Dr. Gerhard Christofori

Prof. Dr. Nicola Aceto

PD Dr. Andrea Banfi

Basel, den 23. April 2019

Prof. Dr. Martin Spiess
Dekan

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

My PhD Thesis is the result of a full commitment and dedication over the last three and a half years. Yet, it is not solely my achievement and I would like to acknowledge many contributors for their enormous involvement.

First and foremost, I would like to thank my direct supervisor, Prof. Nicola Aceto. I fearlessly took the risk and terminated my first PhD project with the confidence that I will receive his full support. He helped me to develop further and to achieve goals that were out of my reach before. I will always be grateful for his enthusiasm, availability and trust in my capabilities.

This work was closely supervised and supported by two PhD Thesis Committee members: Prof. Dr. Gerhard M. Christofori and PD Dr. Andrea Banfi. I would like to express my deep gratitude for their suggestions and assistance. I am also very grateful to Prof. Dr. Niko Bereenwinkel for project inputs that he provided over the course of my studies.

Establishing a close collaboration with clinicians from Basel was not only vital for my PhD Thesis, but also provided a strong, patient-based drive for my research. To all, clinicians and their patients who agreed to take part in the 'CTC Study', but especially to Marcus, Julia, Mathilde, Walter and Christoph – I cannot thank you enough!

I would also like to thank three incredible friends with whom we have placed our first steps as Cancer Metastasis lab (CMET): Ilona, Cinzia and Sofia. Thank you for being part of my journey and making everyday lab life a real pleasure.

Special thanks to another CMET member, Francesc, for his involvement in all of my projects and enormous patience.

All members of the CMET lab deserve a huge thanks for their support on both professional and personal level, especially Massimo who kindly provided suggestions for this PhD Thesis.

CMET group was strongly supported by the Tumor Biology lab and all its members deserve acknowledgements for their contribution. Special thanks to Steffi, Fabiana, Meera, Maren and Nami.

I am also extremely grateful to all my previous supervisors, Prof. Radek Skoda, Prof. Umesh Deshmukh, Prof. Harini Bagavant and Prof. Agnieszka Szalewska-Pałasz for their enormous impact on my education. Thank you for your continuous support!

My experience as a PhD student would not be the same without two years spent in the Experimental Hematology lab. Big thanks to my dear 'frolleagues': Taka, Yukiko, Morgane, Ronny, Gabi, Helene, Lucia, Hui, Jakub x2, Pontus, Axel, Renate and others.

Huge thanks also to David who made my time at ZLF very enjoyable.

Performing all necessary experiments would not be possible without people from Core Facilities. Many thanks to Emilka, Angelika, Emmanuel, Robert, Pascal, Christian, Katja and Elodie!

Last but not least, I would like to thank people that indirectly contributed to this work and were the source of my daily motivation. My beloved partner - Yves, whose understanding, devotion and faith in me made me a better person. My whole family, but especially my Sister, closest friend and my role model – Kasia. And of course, all my friends, especially those that enjoy playing board games with me. Thank you!!!

THESIS STATEMENT

This PhD Thesis has been performed in the Cancer Metastasis lab under the supervision of Prof. Nicola Aceto from the Department of Biomedicine, at the University of Basel and the University Hospital Basel. Presented work started in November 2015 and was completed in early 2019.

Collected results were structured as a cumulative dissertation and consist of a general introduction relevant for my work, followed by applied methods and results sections, formed predominantly of first-author manuscripts. This data was brought together with unpublished observations gathered during my studies. Finally, I provide conclusions were key aspects of the performed research and proposed future plans are discussed.

Notably, all computational analyses were performed by Dr. Francesc Castro-Giner.

ABBREVIATIONS

AF488: Alexa Fluor 488

CA 15-3: Cancer antigen 15-3

CBX6: Chromobox 6

CDK4: Cyclin-dependent kinase 4

CDK6: Cyclin-dependent kinase 6

CRP: C-reactive protein

CTC: Circulating tumor cell

EDTA: Ethylenediaminetetraacetic acid

EGFR: Epidermal growth factor receptor

EpCAM: Epithelial cell adhesion molecule

ER: Estrogen receptor

E&T-Seq: Exome and transcriptome sequencing

FITC: Fluorescein isothiocyanate

G-CSF: Granulocyte colony-stimulating factor

GM-CSF: Granulocyte-macrophage colony-stimulating factor

G&T-Seq: Genome and transcriptome sequencing

HCT: Hematocrit

HER2: Human epidermal growth factor receptor 2

HGB: Hemoglobin

LDH: Lactate dehydrogenase

LUC: Large unstained cells

MCH: Mean corpuscular hemoglobin

MCHC: Mean corpuscular hemoglobin concentration

MCV: Mean corpuscular volume

MDA: Multiple displacement amplification

MDSC: Myeloid derived suppressor cell

MPV: Mean platelet volume

mTOR: Mechanistic target of rapamycin

MUC-1: Mucin 1

NK: Natural killer

NO: Nitric oxide

NSG: NOD scid gamma

PD-L1: Programmed death-ligand 1

PLT: Platelets

PR: Progesterone receptor

RANK: Receptor activator of nuclear factor κB

RANKL: Receptor activator of nuclear factor κB ligand

RBC: Red blood cells

ROS: Reactive oxygen species

WBC: White blood cells

WGA: Whole genome amplification

VEGF: Vascular endothelial growth factor

INDEX OF FIGURES and TABLES

INTRODUCTION GRAPHS

FIRST-AUTHOR MANUSRIPTS

"NEUTROPHILS ESCORT CTCS TO ENABLE CELL CYCLE PROGRESSION"

UNPUBLISHED RESULTS

1. THESIS SUMMARY

Metastatic dissemination is responsible for more than 90% of cancer-related deaths. However, the molecular features underlying the metastatic process are largely uncharacterized. Cancer cells that leave a primary tumor and enter the bloodstream are referred to as circulating tumor cells (CTCs). While extraordinarily rare compared to normal blood cells, their isolation and characterization offers a unique opportunity to study how metastasis occurs. Recent breakthrough developments in microfluidics technology have enabled a more detailed examination of blood samples, highlighting that tumor cells in circulation are organized as single and clustered CTCs, with the latter being associated with a greater metastatic potential. Moreover, an additional subtype of CTCs was observed and characterized as white blood cells (WBCs) – associated CTCs with unknown prevalence and function.

The goal of my PhD Thesis was to investigate CTCs in depth by implementing a novel approach in cancer metastasis research. For the first time, a single cell-resolution, parallel genome and transcriptome sequencing of single, clustered and WBC-associated CTCs of cancer patients and metastatic mouse models has been applied. This technique provided a mean for multiple observations including a breakthrough discovery – WBCs circulating in association with tumor cells have a crucial impact on their metastatic potential and these cell aggregates were described as the most metastatic CTC subtype. Moreover, single-cell RNA sequencing revealed a specific pattern of these WBCs, with neutrophils representing the majority of the cases. Additionally, cell-cell junction and cytokine-receptor pairs that define CTC-neutrophil clusters were described and proposed as key vulnerabilities of the metastatic process. Thus, the association between neutrophils and CTCs fuels cell cycle progression within the bloodstream and expands the metastatic potential of CTCs, providing a rationale for targeting this interaction in breast cancer.

Over the course of three and a half years more than 160 patients with diverse cancer types agreed to donate blood for the ‘CTC Study’. The association between crucial clinical parameters and CTC characteristics have been evaluated in a selected group of 73 consecutive patients characterized by progressive invasive breast cancer, high tumor load and treatment discontinuation. Among these patients, a correlation between a treatment with the monoclonal

antibody denosumab and the absence of CTCs was observed. Additionally, low red blood cell count was associated with the presence of CTCs, whereas high CA 15-3 tumor marker, high mean corpuscular volume, high white blood cell count and high mean platelet volume associated specifically with CTC clusters. These findings carry clinical applications, however further studies will be needed to validate the involvement of denosumab in the prevention of CTC generation.

2. INTRODUCTION

2.1. CANCER METASTASIS

Cancer is a commonly known group of diseases occurring as a result of uninhibited cell growth in various organs. The likelihood of certain tissues to give raise to cancer is highly correlated with the total number of cell divisions necessary for homeostasis [1]. Majority of cases (i.e. solid cancers) are presented with abnormal cell formations called neoplasms (from greek *neo* ("new") and *plasma* ("formation") or simply tumors. Globally, around 15 million new cases and 8 million deaths occur every year, with the frightening prediction of 27 million new patients to be diagnosed in 2030 [2, 3]. Notably, the development of a metastatic disease accounts for more than 90% of cancer-related deaths and clinical data indicates a strong correlation between metastases occurrence and poor prognosis [4].

Metastasis has been described as a series of sequential steps that begins with the intravasation of primary cancer cells into the bloodstream, followed by the survival within circulation, extravasation into secondary distant sites and finally outgrowth of metastatic cancer cells deposits [5]. Initiation of this cascade can result from a cellular stress upon nutrient deprivation and it requires acquisition of features (e.g. via mutagenesis) that allow cells to separate from the tumor mass and avoid anoikis, a detachment-induced apoptosis [6, 7]. This phenomenon is observed in patients when necrotic regions are more common in higher grade (i.e. more aggressive) tumors and they positively correlate with metastatic occurrence [8, 9]. Another crucial factor enabling metastatic processes is the ability to evade immune attacks and to “trick” other cells to serve as support to tumor growth and invasiveness [10]. Interestingly, certain cancers tend to spread only to specific organs. For example, breast cancer often migrates to multiple tissues thorough the body, including the brain, bones, lungs and liver, whereas colon cancer favors seeding mostly to the liver [11, 12]. Detailed evidence explaining differences in frequency or site-specific secondary lesions across different tumor types is yet to be described.

Asymptomatic early-stage cancers tend to be difficult to detected without proper screening programs and the dissemination of cancer cells may have already occurred at the time of diagnosis (e.g. lung cancer) [13]. Initially, these scattered tumor cells are undetectable and form only small colonies (micrometastases). Studies have shown that disseminated tumor cells can survive in dormancy even 20 years before they will cause a secondary disease outbreak [14].

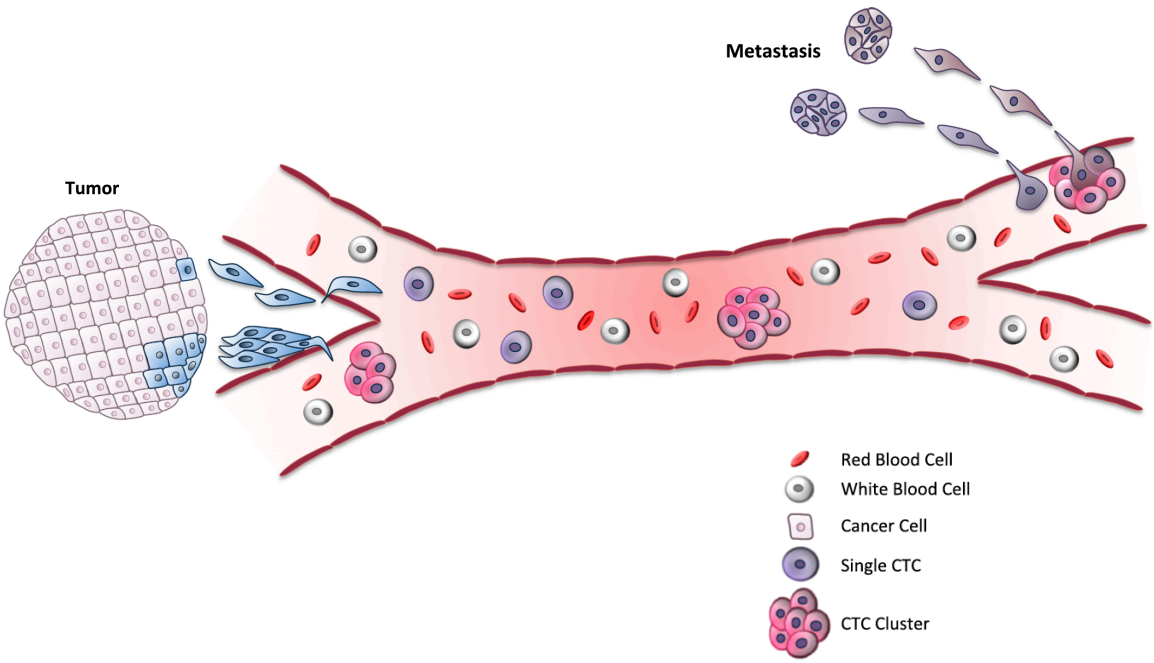
Currently, patients that developed a metastatic disease are considered incurable. One of the main reasons for this outcome is treatment difficulty due to the intra- and inter-patient heterogeneity. No two cases are identical and therefore clinical choices are often problematic. Moreover, cancer is a dynamic disease with genetic and epigenetic changes occurring as a consequence of treatment-induced selection [15]. Many metastatic patients struggle with a recurring disease (i.e. cancer relapse) and therapy-resistance, therefore consecutive lines of treatment are administered until a non-responsive terminal cancer develops [16]. This explains ongoing efforts to develop relapse prediction tools based on the molecular variables of primary tumor [17]. Additional obstacle is the difficulty of direct sampling and characterization of metastatic lesions, which reflects into poor treatment strategies [18]. It is in this context that analysis of circulating tumor cells (CTCs) on their way to form a secondary lesion may offer an extraordinary opportunity to characterize the molecular features of a progressing metastatic disease, leading to personalized treatment [19].

2.2. BIOLOGY AND CLINICAL SIGNIFICANCE OF CTCS

CTCs originate from a solid tumor, which sheds cells in order to spread. First description of these phenomenon is dated back to the 19th century, when an autopsy of a metastatic cancer patient revealed similar cells in the blood and distant tumors. This observation suggested that cancer cells could be carried in the circulation to reach secondary sites. Since then, the vital role of CTCs in metastatic processes has been suspected, however detailed studies were not possible due to the isolation difficulties. It is estimated that one CTC can be surrounded by ten million of white blood cells and ten billion of red blood cells causing the analysis extremely challenging. Recent development of novel isolation technologies enabled first attempts in characterization of these unstudied cells.

Tumors can spread via single CTCs or CTC clusters [20], however the reasons for these subtypes to co-exist are still far from being completely understood. A more recent study described that CTC clusters are kept together by intraepithelial cell junctions and are up to 50 times more metastatic than single CTCs [21]. Furthermore, CTC clusters have occasionally been found associated to platelets [21], leukocytes [22, 23], as well as to sporadically express mesenchymal markers [24]. Interestingly, clustering with white blood cells (WBC) has been shown to correlate with worse prognosis [25]. Altogether, these characteristics contribute to the high metastatic potential of CTC clusters [26], yet they do represent points of attack for subverting a metastatic disease.

In epithelial cancers (e.g. breast or lung), primary tumors and metastatic deposits generally display an epithelial phenotype. In contrast, CTCs have been shown to persist in a dynamic state while associated to platelets and leukocytes, and to display both mesenchymal and epithelial phenotypes [24]. Several studies have suggested that the epithelial to mesenchymal-transition (EMT) contributes to early-stage dissemination of cancer and is pivotal for invasion and metastasis. However, recent results seem to oppose this model, showing that mesenchymal phenotype is associated with therapy resistance, but not required for tumor dissemination [27, 28].



Introduction Figure 1. Schematic view of metastatic formation by CTCs. Tumor cells intravasate as either single or clustered cells and are carried in the bloodstream toward the secondary sites. Upon extravasation a metastatic outgrowth may occur. (adapted from Gkountela *et.al.,* 2016 [29])

CTCs have been proven beneficial in prediction and monitoring of treatment responses. For example, detection of mutation changes in CTCs were shown to reflect the therapy effects in lung cancer [30]. Additionally, molecular analysis of blood-derived tumor cells (e.g. expression of HER2, hormone receptors ER and PR) can be vital when tumor biopsy is impossible due to localization. This approach has been described as liquid biopsy (i.e. blood-derived) and is appreciated as a non-invasive method of cancer screening. Moreover, some biomarkers can change during disease recurrence and evaluation of CTC status can be used for making clinical decisions.

Enumeration of CTCs can also carry an evident prognostic value. In a recent retrospective analysis of 2436 metastatic breast cancer patients, a threshold of five or more CTCs was used to determine the group of patients with a shorter overall survival [31]. Results were indicative for all studied patients, irrespective of the breast cancer subtype (i.e. hormone receptor-positive, HER2-positive or triple-negative) and regardless of *de novo* or recurrent metastatic disease. Altogether, reliable CTC detection methods are currently highly desirable to facilitate personalized treatment choices.

2.3. CTCS DETECTION AND ISOLATION METHODS

The fact that CTCs are surrounded by billions of blood cells has hindered their analysis until recently. In order to be effective, the applied detection method needs to be specific and sensitive enough to capture these rare cells. As a potential clinical tool, it also requires high reliability and cost-efficiency necessary for routine analyses. An additional consideration while choosing the appropriate procedure includes the maintenance of cell viability, which is required for certain downstream analyses (e.g. culture or drug-testing).

First-generation detection methods of CTCs were based on biological features of CTCs, such as the expression of epithelial-specific markers (e.g. EpCAM or cytokeratin), which are absent in the normal blood cells [32]. An additional staining detecting white blood cells (anti-CD45) may be applied to increase the detection confidence. This approach is utilized in CellSearch technology [33], the only FDA-approved tool for CTC enumeration from the blood of patients with cancer, which uses magnetic beads coated with anti-EpCAM antibodies to capture CTCs. Another example of immunoaffinity based CTC enrichment platform is 'CTC-chip' with anti-EpCAM antibodies covering a large surface of microfluidic chip, that enabled identification of CTCs in 115 of 116 studied metastatic cancer patients [34]. Importantly, the positive selection and antigen-dependent approach can overlook cells that express low levels of the markers of choice. Therefore, antigen independent methods focused on depletion of red and white blood cells (iChip [35]), physical aspects (Cluster Chip [23], Parsortix [36]), centrifugation [37] or ultrasound-based enrichment [38] have emerged. Until now, CTC cluster-specific methods are represented by the minority of available methods.

Varying with the isolation method of choice, molecular characterization of CTCs can now be applied upon further CTC micromanipulation and purification. To this end, technologies that facilitate CTC enrichment continue to emerge and enable a full spectrum of downstream applications including DNA and RNA sequencing, CNV characterizations and others.

CTC isolation and characterization methods are providing fundamental insights into those mechanisms that underlie the metastatic process. Unquestionable influence on these complex biochemical and biological changes is attributed to immune cells in tumor microenvironment [39-42].

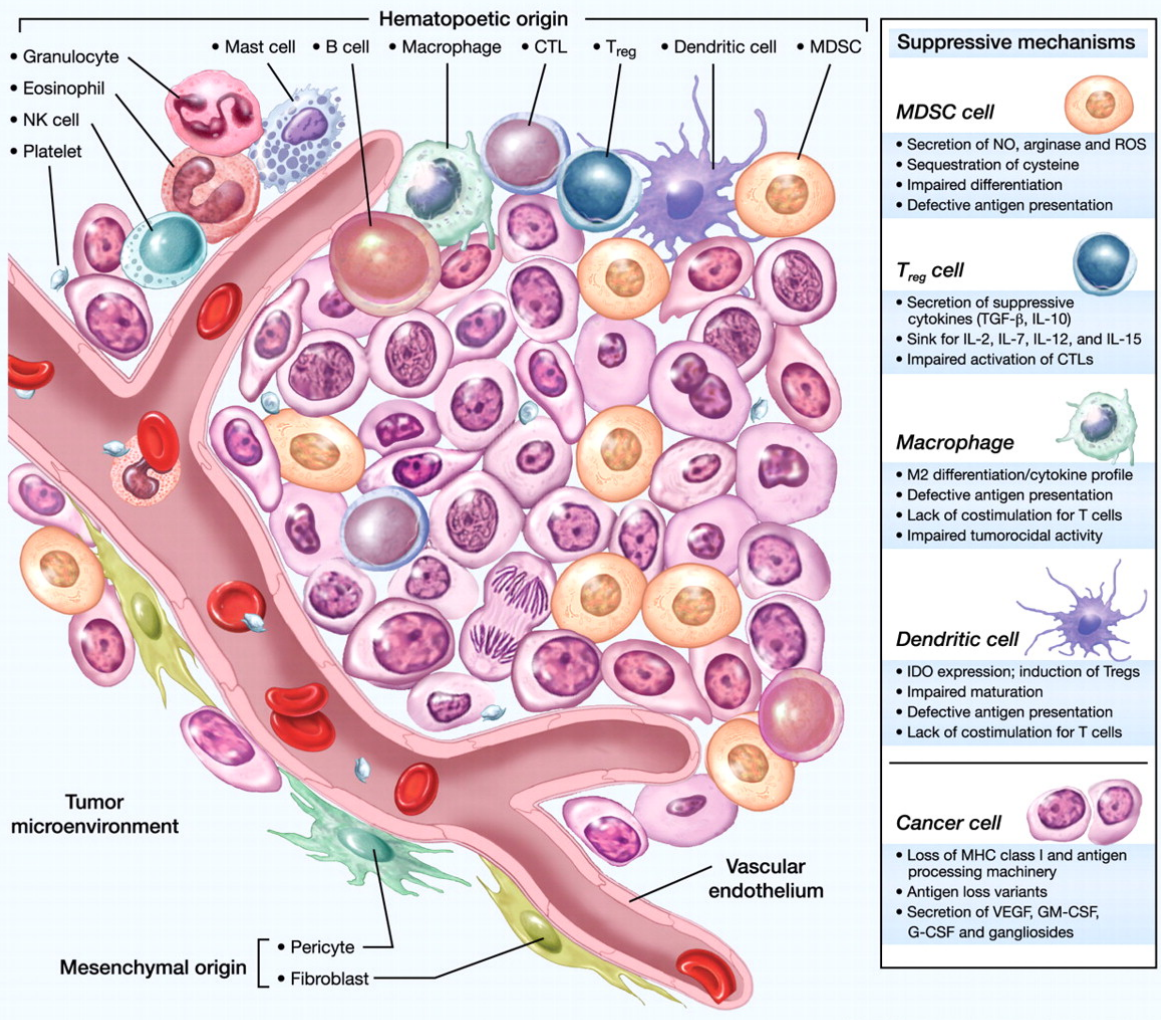
2.4. IMMUNE REGULATION OF METASTASIS

Principles of an immune-dependent tumor progression can be divided into three steps: elimination, equilibrium and escape. The first stage involves an initial contact between neoplastic and immune cells, which typically results in eradication of most or all tumor cells. The second stage describes the process during which both sides start to co-exist, while further selective pressure and clonal selection of neoplastic cells occurs. Lastly, the third phase is characterized by an active growth of a tumor, regardless of the opposed efforts from the immune system (i.e. immune escape). Mounting evidence indicates the tumor *intelligence* seen as turning of the hostile immune system into co-conspirator and conversion of other cell types into potent immunosuppressive agents.

Multiple non-malignant cells co-exist with cancer cells forming together a tumor niche, including neutrophils [43], myeloid-derived suppressor cells MDSCs [44], T cells [45], macrophages [46] and dendritic cells [47]. These cells are not just bystanders as both anti- and pro-tumor activity involving them have been described [48-51]. In this respect, metastasis promotion can take on multiple forms. Firstly, immunosuppressive activity protects from surveillance by killer cells, such as MDSCs inhibition of hostile $CD8^+$ T cells [52] or PD-L1-mediated checkpoint inhibition of natural killer (NK) cells by tumor-associated macrophages [53]. Then, maintenance of a chronic inflammatory environment by infiltrating immune cells leading to further recruitment of pro-tumor immune cells leading to promotion of cancer spread [54]. Additionally, secretion of vascular endothelial growth factor (VEGF) by metastasis-associated macrophages attracts vasculature into the tumor and induces vascular permeability, which facilitates extravasation of cancer cells [55].

CTCs after leaving the tumor site might become vulnerable to immune-associated elimination. In the context of liquid biopsy, a positive correlation has been observed between the presence of CTCs and immune cells in the circulation (i.e. MDSCs and T regulatory cells) [56, 57]. Considering that also increased circulating levels of pro-inflammatory IL-1β and IL-6 are strongly predictive of poor clinical outcome, it seems logical that tumor spread might be supported by the suppression of peripheral antitumor responses. Further studies are required to address the question whether the immunosuppressive mechanism is tumor-specific or whether

they represent the systemic immunomodulation. Moreover, a causal relationship between CTCs and pro-tumor circulating cells needs an additional clarification.



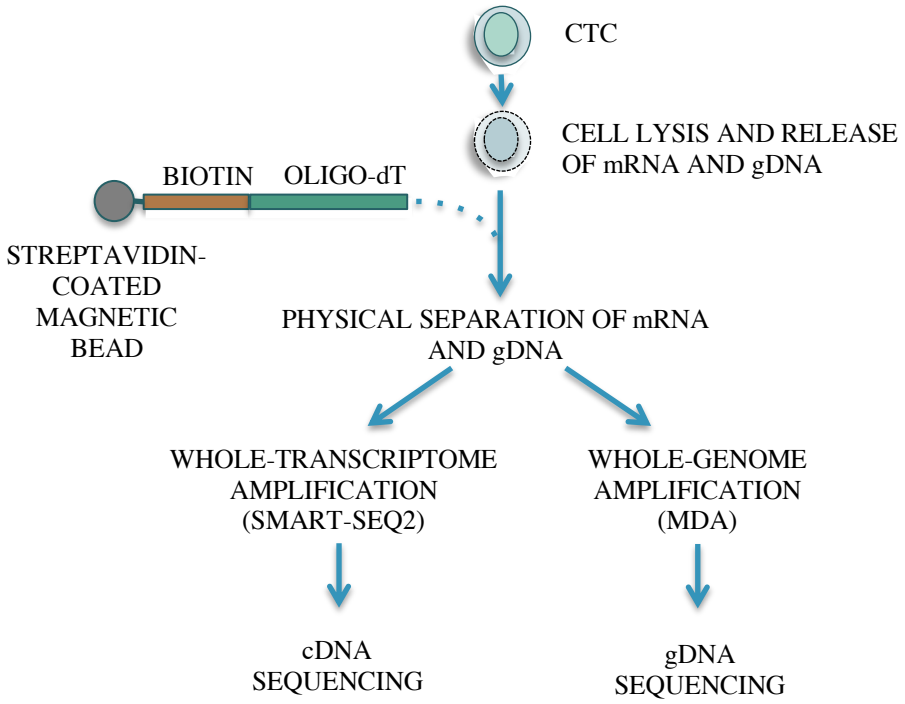
Introduction Figure 2. Tumor niche with infiltrating immune cells. Recruited non-malignant cells play vital role in regulatory networks and promote tumor expansion. G-CSF, granulocyte colony-stimulating factor; GM-CSF, granulocyte-macrophage colony-stimulating factor; NK, natural killer; NO, nitric oxide; ROS, reactive oxygen species. (adapted from Kerkar *et.al.,* 2012 [58]).

Detailed characterization of immunomodulatory mechanisms in cancer has led to development of novel therapeutic strategies ([59-61]). Importantly, the idea of targeting immune-checkpoint inhibitors in order to restore the antitumor immunity is currently the main focus of numerous ongoing clinical trials.

2.5. SINGLE-CELL GENOME AND TRANSCRIPTOME SEQUENCING

Single cell-resolution approaches have been successfully implemented for genome or transcriptome analysis [62-64], providing a detailed characterization of cancer cells and assessment of their heterogeneity. Collected expression patterns, mutational profiles and copy-number analyses can be utilized in the clinic as both treatment evaluation and prognostic tools.

Until recently, technical challenges have held research back from understanding transcriptome dynamics in relation to the mutational profile of each single tumor cell. For example, parallel genome and transcriptome sequencing (G&T-Seq) method [65] offers a unique tool for such analysis, which could not previously be obtained from DNA or RNA sequencing alone.



Introduction Figure 3. Parallel Genome and Transcriptome Sequencing Workflow. Each single circulating tumor cell (CTC) will be lysed and mixed with beads capturing polyA mRNA. After magnetic separation, both mRNA and gDNA will be amplified using Smart-seq2 or Multiple Displacement Amplification (MDA), respectively, followed by sequencing (based on Macaulay et.al.[65])

Single-cell multiomics can also include epigenome [66] and protein [67] analyses. Undoubtedly, further interrogations of multiple parameters in parallel can be extremely challenging (e.g. interpretation of noisy data), but also highly beneficial, not only in the context of cancer research.

3. METHODS

3.1. PATIENT SELECTION

After obtaining written informed consent, breast cancer patients donated 7.5 –15ml of blood in EDTA vacutainers at least once. Involved patients were characterized by invasive breast cancer, high tumor load, progressive disease at the time of blood sampling and treatment discontinuation at the time of CTC isolation (before the next line of therapy). Preselection for breast cancer subtype or specific metastatic sites was not applied during enrollment for the study. Eligible patients were > 18 years old with any menopausal status and had an Eastern Cooperative Oncology Group performance status of 0–3. Disease had to be measurable by Response Evaluation Criteria in Solid Tumors (RECIST) version 1.1 or non-measurable bone only disease. Tumor load was defined by either the size of the primary tumor or the number and size of metastatic lymph nodes or distant sites, and patients with higher tumor load were prioritized. All blood specimens were obtained at the University Hospital Basel under the study protocols approved by the Swiss authorities (EKNZ, Ethics Committee northwest/central Switzerland) and in compliance with the Declaration of Helsinki.

3.2. CLINICAL PARAMETER ASSESSMENT

Primary tumor samples were collected at the initial diagnosis, and IHC was performed for estrogen receptor (ER), progesterone receptor (PR), HER2 and Ki-67. If the patient had primary metastatic disease, a biopsy from the metastatic site was obtained when possible, including marker assessment: ER, PR and HER2. Histopathological diagnosis was conducted by two independent pathologists from the breast cancer unit at the University Hospital Basel. All patients were treated at the Breast Cancer Unit University Hospital Basel according to local standard operating procedures and National Comprehensive Cancer Network and European Society for Medical Oncology guidelines by senior breast oncologists. If a patient had a progression within new distant sites, a new biopsy from that site was taken, when possible, to determine ER, PR and HER2. Patients under systemic treatment had tumor assessment at least every 12 weeks with computed tomographic scans or earlier if tumor progression was anticipated. CTC collection was performed at progression and prior to the next line of therapy or before any treatment was conducted. The patients' data was retrieved by detailed retrospective chart review. Data collection included demographics and disease-specific and treatment-specific data including age, gender, primary stage, histologic subtype, ER/PR/HER2 status, grading, Ki-

67, date of primary diagnosis and relapse, type of relapse (localized, metastatic), site of distant disease, bone-modifying agents (bisphosphonates, denosumab), palliative irradiation, and type of systemic treatment, including time on treatment and time to next subsequent treatment. Data was correlated with CTC counts.

3.3. BLOOD PARAMETER ASSESSMENT

Complete blood counts were measured with the ADVIA 120 Hematology Analyzer (Siemens Healthcare Diagnostics, Tarrytown, NY, USA) using Multispecies version 5.9.0-MS software (Bayer Diagnostics, Tarrytown, NY, USA). Blood samples were taken before each new therapy cycle or at least every month, including cancer antigen 15-3 (CA 15-3), alkaline phosphatase, Ca2+, C-reactive protein, lactate dehydrogenase, red blood cells (RBC), hemoglobin, hematocrit, mean corpuscular volume (MCV), mean corpuscular hemoglobin concentration, white blood cells (WBC), neutrophils, lymphocytes, monocytes, eosinophils, basophils, large unstained cells, platelets and mean platelet volume (MPV). In the vast majority of cases, blood samples were taken simultaneously with the CTC sample or within 7 days after CTCs were taken. Eight of 73 patients had only partial data available, whereas no blood counts were reported at the time of CTC detection for nine of 73 patients.

3.4. CELL CULTURE

MDA-MB-231 LM2 human breast cancer cells (obtained from Dr. Joan Massagué, MSKCC, NY, USA) and 4T1 murine breast cancer cells (ATCC) were grown in DMEM medium (#11330-057, Gibco) supplemented with 10% FBS (#10500064, Gibco) and antibiotic/antimycotic (#15240062, Gibco) in a humidified incubator at 37 °C with 20% O_2 and 5% CO_2. Py2T cells were a gift from Dr. Gerhard Christofori (University of Basel, Switzerland). Human CTC-derived BR16 cells were generated as previously described [68] and propagated as suspension cultures in a humidified incubator at 37 °C with 5% O_2 and 5% CO_2. All cell lines were transduced with lentiviruses carrying GFP-Luciferase (*GFP*) at a multiplicity of infection (MOI) < 5. For CRISPR screening 4T1 cells were transduced with the lentiviruses carrying pLentiCas9-*EGFP* (#63592, Addgene) at a MOI of 1. *GFP*-positive cells were sorted as single cells into 96-well plates and cultured as clonal cell lines. Lines with 100% *GFP*-positivity were kept and Cas9 expression was confirmed by western blotting (#844301, Biolegend).

3.5. MOUSE EXPERIMENTS

All mouse experiments were carried out according to institutional and cantonal guidelines (approved mouse protocol #2781, cantonal veterinary office of Basel-City). Nod Scid Gamma (NSG) and Balb/c mice were purchased from The Jakson Laboratory (Bar Harbor, Maine, USA) and kept in pathogen-free conditions, accordingly to institutional guidelines. Transgenic MMTV-PyMT mice were obtained from Dr. Gerhard Christofori (University of Basel). Orthotopic breast cancer lesions were generated in 8-10 weeks old NSG females upon the injection with either $1x10^6$ LM2-*GFP*, $0.5x10^6$ 4T1-*GFP* or $1x10^6$ BR16-*GFP* cells into the mammary fat pad. Similarly, Balb/c mice received a syngeneic graft of $0.5x10^6$ 4T1-*GFP* cells. In all cases, breast cancer cells were inoculated in 100µl of 50% Cultrex PathClear Reduced Growth Factor Basement Membrane Extract (#3533-010-02, R&D Biosystems) in PBS. Blood draw for CTC analysis and organ dissection were performed after 3 weeks for NSG-4T1-*GFP*, 4-5 weeks for Balb/c-4T1-*GFP* and NSG-LM2-*GFP*, 5 months for NSG-BR16-*GFP* and at 13 weeks of age for MMTV-PyMT mice. Generally, immunocompetent models (Balb/c-4T1-*GFP* and MMTV-PyMT) developed a primary tumor that reached the maximum allowed size before developing overt metastatic disease. For this reason, they were rather used throughout the manuscript as models to assess direct metastatic potential of cancer cells injected directly in the venous circulation (i.e. tail vein). In contrast, immunocompromised models (NSG-4T1-*GFP*, NSG-LM2-*GFP* and NSG-BR16-*GFP* mice) were used as the preferred system to assess spontaneous CTC and metastasis formation from the primary tumor. All experiments whereby both immunocompetent and immunocompromised mice were used side by side have led to the same conclusions.

3.6. CTC CAPTURE

Human CTCs were captured from unprocessed peripheral blood samples with the Parsortix microfluidic device using Cell Separation Cassettes (GEN3D6.5, ANGLE), within 1 hour from blood draw. Next, in-cassette staining was performed with an antibody cocktail containing antibodies against EpCAM-AF488 (#CST5198, Cell Signaling Technology), HER2-AF488 (#324410, BioLegend), EGFR-FITC (#GTX11400, GeneTex) and CD45-BV605 (#304042, BioLegend). For mouse experiments, 500–1000µl of blood was collected through cardiac puncture and processed immediately on the Parsortix microfluidic device. For tumor-draining vessel experiments, the tumor was first exposed by opening the mouse flank. The largest tumor-

associated vessel was then localized and approximately 2µl of blood was collected upon a small incision. CTCs from the MMTV-PyMT mouse model were stained with antibodies against mouse EpCAM-AF488 (#118210, BioLegend) and CD45-BV605 (#103140, BioLegend). For all other models (xenografts and syngeneic), carrying cancer cells stably expressing a GFP Luciferase reporter, only anti-CD45 staining was performed, while CTCs were identified based on GFP expression. The number of captured CTCs, including single CTCs, CTC clusters and CTC-WBC clusters, was determined while cells were still in the cassette. CTCs were then released from the cassette in DPBS (#14190169, Gibco) onto ultra-low attachment plates (#3471-COR, Corning). Representative pictures were taken at 40x magnification with Leica DMI4000 fluorescent microscope using Leica LAS and analyzed with ImageJ.

3.7. ASSESSMENT OF THE DIRECT METASTATIC POTENTIAL OF CTCS

8-10 weeks old NSG females were injected with $0.5x10^6$ 4T1-*GFP* or $1x10^6$ BR16-*GFP* cells. 8-10 weeks old Balb/c females were injected with $0.5x10^6$ 4T1-*GFP* cells. Upon tumor development, blood was collected via heart puncture and run through the Parsortix device. Single CTCs, CTC clusters and CTC-neutrophil clusters were individually micromanipulated and 100 cells per mouse (for NSG-4T1-*GFP* model) or 500 cells per mouse (for NSG-BR16-*GFP* and Balb/c-4T1-*GFP* models) from each category were injected into the tail vein of recipient mice (NSG or Balb/c, respectively). Metastasis onset and growth rate in lungs was noninvasively monitored on a weekly schedule with the IVIS system, or through lung immunohistochemical staining of pan-cytokeratin (#GTX27753, Genetex) at the time of experiment termination.

3.8. WHITE BLOOD CELL SORTING

Reference WBCs were obtained from the peripheral blood of breast cancer patients (n=5) and healthy individuals (n=3) after signing informed consent, naïve NSG and Balb/c mice (females at 8-12 weeks), Balb/c-4T1-*GFP* and NSG-CDX-BR16-*GFP* mouse models at the time of experiment termination. In brief, red blood cells, granulocytes and mononuclear cells were separated by gradient centrifugation with Lymphoprep (#1114545, STEMCELL Technologies). Desired fractions were manually isolated and washed with 1% BSA/PBS buffer. Additionally, the granulocyte fraction was purified from contaminating red blood cells by 10 minutes incubation in 0.16M ammonium chloride. Unspecific antibody binding was prevented by blocking the Fc receptor for 15 minutes (human: #422301, BioLegend; mouse: #101320,

BioLegend). Cells were stained with white blood cell markers: human – anti-CD14-APC (#301808, BioLegend), anti-CD66b-FITC (#305104, BioLegend), anti-CD3-BV421 (#317344, BioLegend), anti-CD19-FITC (#302206, BioLegend), anti-CD335-PE (#331908, BioLegend) anti-CD41-PE/Cy5 (#303708, BioLegend); mouse – anti-Gr-1-APC/Cy7 (#108423, BioLegend), anti-CD11b-APC (#101211, BioLegend), anti-CD3-BV421 (#100227, BioLegend), anti-CD19-FITC (#115505, BioLegend) or anti-CD19-BV605 (#115539, BioLegend; for mouse models with GFP reporter), anti-CD49b-PE (#108907, BioLegend), anti-CD41-PE/Cy5 (#133921, BioLegend). Cell populations were determined by the expression of characteristic markers: for human granulocytes ($CD66b^+CD41^-$), monocytes ($CD14^+CD3^-CD19^-CD335^-CD41^-$), T cells ($CD14^-CD3^+CD19^-CD335^-CD41^-$), B cells ($CD14^-CD3^-CD19^+CD335^-CD41^-$), NK cells ($CD14^-CD3^-CD19^-CD335^+CD41^-$); for mouse granulocytes ($Gr\text{-}1^+CD41^-$), monocytes ($CD11b^+CD3^-CD19^-CD49b^-CD41^-$), T cells ($CD11b^-CD3^+CD19^-CD49b^-CD41^-$), B cells ($CD11b^-CD3^-CD19^+CD49b^-CD41^-$), NK cells ($CD11b^-CD3^-CD19^-CD49b^+CD41^-$). One hundred cells from each population were sorted (FACSAria III, BD Biosciences) directly into microcentrifuge tubes containing 2.5µl RLT Plus lysis buffer (#1053393, Qiagen).

3.9. NEUTROPHIL CO-CULTURE WITH TUMOR CELLS

Human neutrophils were purified from healthy donor blood upon gradient centrifugation with Lymphoprep™ (Stemcell Technologies). 8'000 neutrophils were added to 100'000 LM2, BR16 or Brx50 CTC-derived cells and co-cultured in CTC media for 72 hours. Then, gDNA was isolated from tumor cells (or untreated control cells) and processed for whole exome sequencing.

3.10. EXOME AND TRANSCRIPTOME SEQUENCING

Individual cells from CTCs alone or CTC-WBC clusters were mechanically separated with gentle micromanipulation (CellCelector, ALS). AF488/FITC-positive (or GFP-positive) and BV605-negative CTCs or AF488/FITC-negative and BV605-positive WBCs were immediately transferred into individual tubes (#321-032-501, Axygen) containing 2.5µl RLT Plus lysis buffer (#1053393, Qiagen) and 1U/µl SUPERase In RNase Inhibitor (#AM2694, Invitrogen). Samples were immediately frozen on dry ice and kept at -80°C until further processing. Following previously published protocol for parallel DNA and RNA sequencing from individual cells [69], genomes and transcriptomes of lysed cells were separated and amplified (#25-6601-97, GE Healthcare for genome and Smart-seq2 from for transcriptome). Reference white blood

cells were prepared solely with Smart-seq2 protocol. Libraries were prepared with Nextera XT (Illumina), exomes were enriched using SureSelect XT Human All Exon v6+Cosmic kit (Agilent technologies) and sequenced on HiSeq 2500 (Illumina) in 100bp paired-end mode for DNA sequencing and on NextSeq 500 (Illumina) 75bp single read mode for RNA sequencing.

3.11. DIFFERENTIAL WHITE BLOOD CELL STAINING ON CTC-WBC CLUSTERS

Live CTCs captured within the Parsortix microfluidic cassette were stained with anti-Biotin-CD45 (#103104, BioLegend) and detected with Streptavidin-BV421 (#405226, BioLegend), anti-mouse Ly-6G-AF594 (#127636, BioLegend) and anti-CD11b-AF647 (clone M1/70, kind gift from Dr. Roxane Tussiwand, University of Basel) or with anti-F4/80-AF594 (#123140, BioLegend) and CD11b-AF647. Additionally, MMTV-PyMT-derived CTCs were marked with EpCAM-AF488 (#118210, BioLegend). Next, cells were gently released from the microfluidic system into ultra-low attachment plate and immediately imaged (Leica DMI400). The number of CTC-WBC clusters with neutrophils (Ly-6G^{+}CD11b^{med}), monocytes (Ly-6G^{-}CD11b$^{med/high}$) and macrophages (F4/80^{+}CD11b^{+}) was assessed. Immediately after imaging, cells were centrifuged (500rpm, 3 minutes) on a glass slide and fixed in methanol for 1 minute. After brief air-drying, slides were stained using Wright-Giemsa stain kit (#9990710, ThermoFisher) to visualize nuclear morphology of captured cells, following the manufacturer's instructions.

3.12. IMMUNOFLUORESCENCE STAINING

Formalin-fixed, paraffin-embedded (FFPE) sections were obtained from primary tumors and metastatic sites of patients with ER/PR-positive breast cancer (Department of Pathology, University Hospital Basel) who had detectable CTC-WBC clusters. Similarly, mouse-derived primary tumors and metastases were fixed in 4% paraformaldehyde and prepared according to a standard paraffin embedding protocol. Human and mouse sections were handled according to a standard immunofluorescent paraffin-embedded tissue staining protocol. Briefly, after deparaffinization in xylene and re-hydratation, antigen retrieval was carried out in 10mM sodium citrate (pH 6.0) at 95°C for 25 minutes. For CTC and DTC staining, cell suspension was centrifuged (3min, 500 rpm) on a coated glass slide (#5991056, ThermoFisher) and air-dried. Cells were fixed in 4% paraformaldehyde for 12 min and stored in PBS until needed. For both FFPE sections and cells, after 1 hour of blocking with 10% horse serum, specimens were co-stained for pan-cytokeratin (#GTX27753, Genetex) detected with anti-mouse IgG-AF488 (#A-

21202, ThermoFisher), myeloperoxidase (#AF3667-SP, R&D) detected with anti-goat IgG-AF568 (A-11057, ThermoFisher), Ki67 (#ab15580, Abcam) detected with anti-rabbit IgG-AF647 (A-31573, ThermoFisher) and DAPI (#D1306, ThermoFisher).

3.13. *IN VITRO* CYTOKINE TREATMENT

100'000 4T1-*GFP* cells per well were seeded in a 6-well plate and cultured in growth medium overnight. Next morning, cells were washed 3 times with PBS and given starvation medium (0.1% FBS). After 48h, the medium was supplemented with 25ng/ml recombinant mouse IL6 (#575702, BioLegend), IL1β (#575102, BioLegend), TNFα (#575202, BioLegend) and OSM (#762802, BioLegend), either individually or in combination. Cells were stimulated for 24h and then harvested upon trypsinization, enumerated using automatic cell counter and 300'000 cells were injected intravenously into 8-10 weeks old female mice.

3.14. MUTAGENESIS

Lentiviral vectors with human *MERTK* (CCSB-Broad LentiORF, CloneId: ccsbBroad304_11503, Dharmacon) and human *TLE1* (Precision LentiORF, CloneId: PLOHS_100005903, Dharmacon) served as base for introduction of specific mutations using QuikChange II XL site-directed mutagenesis kit (#200522, Agilent Technologies). Lentiviral particles were then prepared with Dharmacon Transduction Starter Kit and upon transduction, 4T1-*GFP* cells were selected with 9 µg/ml Blasticidin S for 6 days.

3.15. MYLEOID CELLS DEPLETION

For neutrophil depletion studies in primary tumor models, mice were injected intraperitoneally with 400µg of anti-Ly-6G IgG2a (#127650, BioLegend) or control IgG2a (#400566, BioLegend) when tumors were palpable (day two after injection of 4T1-*GFP* cells, day six after injection of LM2-*GFP* cells and day 30 after injection of BR16-*GFP* cells). Efficiency of immune cell depletion was monitored after 48 hours with Advia120 Hematology Analyzer (Siemens) using Multispecies version 5.9.0-MS software (Bayer). Additionally, NSG-4T1-*GFP* mice received a second dosage of anti-Ly-6G or control IgG2a antibodies (100 µg) on day 19, NSG-LM2-*GFP* mice on day 25 and NSG-BR16-*GFP* mice on day 45. Tumor size was determined with caliper measurements every seven days and tumor volume was calculated using modified ellipsoid formula: $V=1/2(\text{Length} \times \text{Width}^2)$. At termination, lung metastases were

measured with IVIS Lumina II (Perkin Elmer) and metastatic index was determined by normalizing the photon/second count of the metastasis with that of the primary tumor. For neutrophil pre-depletion experiments, a single dose of 400µg of anti-Ly-6G IgG2a was injected intraperitoneally 24h before tumor cells intravenous inoculation. Mice were sacrificed in accordance to our approved protocol and the survival data was inferred accordingly.

3.16. G-CSF OVEREXPRESSION

Human *G-CSF* was transduced into 4T1-*GFP*, LM2-*GFP* and BR16-*GFP* cells using the Precision LentiORF (GE Healthcare) system. Construct-positive cells were selected with 9µg/ml Blasticidin S for 4 days (4T1) or 7 days (LM2, BR16). Overexpression of *G-CSF* was confirmed by qPCR using human-specific primers for LM2 and BR16 cells (Forward: 5'GAGTTGGGTCCCACCTTG3', Reverse: 5'TGGAAAGCAGAGGCGAAG3') or primers recognizing both mouse and human transcripts for 4T1 cells (Forward: 5'TGTGCCACCTAC AAGCTGTG3', Reverse: 5'CCATCTGCTGCCAGATGGTGGT3').

3.17. sgRNA MINIPOOL DESIGN, TRANSDUCTION AND *IN VIVO* TRANSPLANTATION

All sgRNAs were designed using the GPP Web Portal (https://portals.broadinstitute.org/gpp/public/analysis-tools/sgrna-design) and sgRNA oligos were synthesized by Microsynth. Each sgRNA was individually cloned into the pLentiGuide-Puro vector (#52963, Addgene). 4T1-Cas9-*GFP* cells were then transduced separately with each individual sgRNA vector at MOI=0.4. Upon seven days of puromycin selection, 4T1-Cas9-*GFP* cells carrying individual sgRNAs were mixed in equal cell numbers, taken for genomic DNA extraction and, at the same time, subcutaneously injected (≥ 1000 cells per sgRNA; 500'000 total cells) into the mammary fad pad of NSG mice.

3.18. sgRNA SEQUENCING

gDNA was extracted from cells at different stages (prior to injection, upon primary tumor growth and from spontaneously formed CTCs) using salt precipitation. The library preparation was carried out using two-step PCR, where the first PCR amplifies a broad region including the sgRNA sequence cassette and the second PCR adds Illumina sequencing adapters to the products from the first PCR, as described previously[70]. Samples were then sequenced on

NextSeq 500 SR75 sequencers. After quality control using FastQC (https://www.bioinformatics.babraham.ac.uk/ projects/fastqc), reads were trimmed using cutadapt (v1.9.1) and aligned to the sgRNA sequences using bowtie2 (v2.2.9), allowing for one mismatch. The normalized counts of each sgRNA were computed by dividing the number of reads for each sgRNA by the library size.

4. COMPUTATIONAL METHODS

Note. All of these crucial for my projects computational analyses were performed by bioinformatician, Dr. Francesc Castro-Giner. Additional support with E&T-sequencing validation was provided by Dr. Robert Ivanek.

4.1. SINGLE-CELL RNA-SEQ DATA PROCESSING

After sequencing, initial quality assessment for RNA-seq data was performed using FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc), FastQ Screen (https://www.bioinformatics.babraham.ac.uk/projects/fastq_screen), and visualized with MultiQC (v0.8). Adaptor sequences, first 9 base pairs and low-quality ends were removed with Trim Galore (v0.4.2, http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/; parameters : --phred33 --length 36 --clip_R1 9). Trimmed reads were aligned to a combined human (GRCh38) and mouse (GRCm38) genome reference using STAR (v 2.5.2a; parameters: --runMode alignReads --genomeLoad LoadAndExit). Quality control of resulting BAM files was performed with RSeQC (v2.6.4). The gene-level expression counts were computed with featureCounts (v1.5.1) using the gene annotations obtained from RefSeq (release 70). Samples with less than 800 features detected (threshold ≥ 1 mapped read) or showing more than 5% of contamination from the other species were removed from further analysis. To normalize gene counts for cell-specific biases, we used size factors computed utilizing the deconvolution implemented in the scran package (v1.6.5) available on R/Bioconductor. After normalization, the effect of technical factors (library size and number of detected features) on variance was evaluated using t-distributed stochastic neighbor embedding (t-SNE) adjusted by patient or mouse model. CTCs showing a substantial contribution of stromal genes and the absence of cancer-associated genes, and CTC-associated WBCs showing no expression of CD45 were removed from the analysis. scRNA-seq data processing, quality control, and visualization was performed with the help the R/Bioconductor package scater (v 1.6.0).

4.2. REFERENCE COMPONENT ANALYSIS

Reference component analysis (RCA) was utilized to identify single cell types using reference transcriptomes. For human samples, the reference bulk transcriptomes were obtained from the Human U133A/GNF1H Gene Atlas and the Primary Cell Atlas (http://biogps.org/), averaging expression levels when multiple replicates were present. Mouse transcriptomes were obtained from the Mouse GNF1M and MOE430 Gene Atlas (http://biogps.org/). The initial gene

selection for the reference transcriptome panel was performed as previously described [71]. An additional filtering of genes was achieved by removing genes specific to CTCs from the human panel and by selecting highly variable genes (HVGs) from the mouse panel. A gene was defined as CTC-specific if its normalized expression (log-counts) relative to the median across the reference WBCs set exceeded 5 in at least 5% of CTC samples. In mouse, only genes that showed high variability in their expression across reference WBCs were included. In order to select HVGs in mouse, gene-specific variance of expression across reference WBCs was estimated using trendVar and decomposed into biological and technical components using decomposeVar from scran package. Highly variable genes were selected on the basis of their biological component (biological variance ≥ 5) and adjusted *P*-value (threshold ≤ 0.05). A total of 5,279 genes were selected for the human reference panel and 655 for the mouse panel. Projection of each sample onto the reference transcriptome was performed as previously described[71], calculating the Pearson correlation between the $\log_{10}$ (FPKM) values of the scRNA-seq samples and the $\log_{10}$ expression values of the global panel using the functions provided by the RCA R package (v1.0; https://github.com/GIS-SP-Group/RCA). For visualization, reference cell types with a low correlation with query samples and non-immune related features were removed. Hierarchical clustering was performed to cluster samples based on their projection values.

4.3. DIFFERENTIAL EXPRESSION AND GENE SET ENRICHMENT ANALYSES

We determined differentially expressed genes by the edgeR likelihood ratio test method (v3.20.1) using the normalized counts with the deconvolution approach and the robust dispersion of estimates options. Gene set over-representation analysis of KEGG pathways in the differentially expressed genes (adjusted p-value threshold ≤ 0.25) was performed with the kegga method implemented in the edgeR R/Bioconductor package (v3.20.1). Enrichment of the KEGG pathways ‘Cell cycle’ (hsa04110) and ‘DNA replication’ (hsa03030) in patient samples was tested with the self-contained rotation gene set test (roast) from the limma R/Bioconductor package (3.34.2) using the msq option as a gene set summary statistic and 5’000 rotations to compute p-values.

4.4. CYTOKINE AND CYTOKINE LIGAND ANALYSIS

A comprehensive collection of cytokines and their receptors was obtained from KEGG pathway 'Cytokine-cytokine receptor interaction' (accession codes hsa04060 and mmu04060 for human and mouse, respectively). Next, human one-to-one orthologous genes for the mouse gene set were obtained from Ensembl (v91) using the biomaRt (v2.34) R/Bioconductor package in order to combine human and mouse datasets. A cytokine-receptor pair was considered to be expressed in a CTC-neutrophil cluster if the cytokine gene in the neutrophil sample and its corresponding receptor in the CTC were expressed at log2 normalized counts per million mapped reads (CPM) ≥ 5. For CTC-neutrophil clusters containing more than one detached CTC, all possible CTC-neutrophil pairs were considered.

4.5. SINGLE-CELL DNA-SEQ DATA PROCESSING

Paired-end reads were aligned to the GRCh38 human or GRCm38 mouse reference genomes using BWA-mem algorithm (v0.7.13; parameters: -M) (https://arxiv.org/abs/1303.3997) and sorted using SAMtools (v1.3.1). Reads were then deduplicated using Picard MarkDuplicates (v2.9.0; http://picard.sourceforge.net/) on a per-sample basis and local realignment was performed using the Genome Analysis Toolkit (GATK) IndelRealigner (v3.7.0) at the sample and donor level to improve alignment accuracy around indels. Quality control and coverage and exome enrichment statistics were generated using FastQC, CollectHsMetrics from Picard suite, and QualiMap (v 2.2.1) and visualized using MultiQC (v0.8).

4.6. SOMATIC MUTATION CALLING AND MUTATION SPECTRUM

Mpileup files were generated with SAMtools (v1.3.1; parameters: -B -q 40) and variants were called using Monovar (v2016-05-14) on all samples from the same donor simultaneously. Resulting variants were annotated using SnpEff on ENSEMBL v86 (www.ensembl.org), dbSNP (build 150), 1000 genomes project (phase 1), and coding mutations from cosmic (v81) using SnpSift (v4.3p). Somatic mutation rates were calculated as the ratio of the number of somatic variants and the number of nucleotides covered in the exome at ≥ 2x. Putative damaging somatic mutations were identified exclusively in donors with matched WBC sequenced using an empirical filtering strategy removing (1) variants present in public databases (dbSNP, 1000 genomes project) at a frequency ≥ 1% or found in 2 or more founders, (2) variants present in at least one reference WBC sample from the same donor, (3) variant loci not covered in reference

WBC samples (threshold ≥ 3 reads), and (4) likely damaging events (truncating, frameshift or splice site variant). VCF processing, downstream filtering, and analysis was performed using the VariantAnnotation and vcfR R/Bioconductor packages. Trinucleotide context of the somatic mutation spectrum was generated and visualized with the SomaticSignatures package (v2.14.0).

4.7. SURVIVAL ANALYSES

Survival analyses were performed using the survival R package (v 2.41-3). Kaplan-Meier curves were generated and Log-Rank test was used to estimate the significance of the difference in survival between groups. For patients, progression-free survival was defined as the period of time between primary tumor diagnosis and first relapse. For NSG-4T1-*GFP* mouse model analysis, death was selected as the endpoint for the analysis and defined as the moment a given animal had to be euthanized according to our mouse protocol guidelines.

4.8. STATISTICAL ANALYSIS OF CLINICAL PARAMETERS

We first screened our data to exclude variables and patients with high content of missing information, as well as observations with implausible values. Cancer therapies were simplified into three main nonexclusive categories (targeted therapy, chemotherapy and hormone therapy). Some patients had undergone multiple lines of therapy. For this reason, we assessed the effects of accumulated therapies and the therapy at CTC evaluation separately. We investigated the association between the different variables of interest and the presence of CTCs using Fisher's exact test for categorical variables, two-sided Wilcoxon rank-sum test for continuous variables (e.g., complete blood counts) and Kruskal-Wallis test for ordinal variables with more than two levels (e.g., stage at diagnosis). For each test, we present the nominal P value. An estimate and 95% CI are also provided for continuous and two-level categorical variables. The estimate corresponds to the OR in Fisher's exact test and to the estimated median of the difference between samples from both groups in the Wilcoxon rank-sum test. To account for potential confounding variables, logistic regression analysis was conducted, adjusting by age at primary diagnosis, tumor stage at diagnosis, tumor grade and histologic subtype. Adjusted P values were calculated following the Benjamini-Hochberg method, combining all tests performed in this work. Associations with an adjusted P value ≤ 0.05 are highlighted in the text. We conducted the data wrangling and statistical analysis in R (version 3.4.0; R Foundation for Statistical Computing, Vienna, Austria).

5. RESULTS

5.1. FIRST-AUTHOR MANUSCRIPTS

5.1.1. "NEUTROPHILS ESCORT CTCS TO ENABLE CELL CYCLE PROGRESSION" [72]

Circulating tumor cells (CTCs) are precursors of metastasis in various solid cancers including breast cancer [73], and are occasionally found in association to white blood cells (WBCs) [26]. The role of CTC-WBC clusters in metastasis development as well as the principles that govern the interplay between CTCs and WBCs during blood-borne metastasis are largely uncharacterized.

We first sought to determine the number and composition of CTC-WBC clusters in breast cancer patients and mouse models. We obtained blood samples from 70 patients with invasive breast cancer that discontinued their treatment due to progressive disease, as well as from five different breast cancer mouse models, and we enriched for CTCs using the Parsortix microfluidic device [74] (Extended Data Fig. 1a-e). Live CTCs were stained for cancer-associated cell surface markers EpCAM, HER2, and EGFR or imaged directly for the expression of GFP, as well as labeled for CD45 to identify WBCs (Fig. 1a and Extended Data Fig. 1f). Among 70 patients, 34 (48.6%) had detectable CTCs, with a mean number of 22 CTCs per 7.5ml of blood (Supplementary Tables 1 and 2). While the majority of CTCs were single (88.0%), we also detected CTC clusters (8.6%) and CTC-WBC clusters (3.4%) (Fig. 1b and Extended Data Fig. 1g,h). Similarly, we observed that CTC-WBC clusters were present in all tested mouse models, comprising those with immunodeficient or immunocompetent background, ranging from 0.05% to 61% of the total CTC population (Fig. 1b and Extended Data Fig. 1i,j). Importantly, CTC abundance and ratios dramatically changed when drawing blood upstream of capillary beds as opposed to more downstream locations, indicating that clustered CTCs are shed early, yet may be trapped in capillaries before reaching the periphery (Extended Data Fig. 1k-n). Thus, CTC-WBC clusters are rare in the peripheral circulation, yet consistently found across breast cancer patients and mouse models.

We then asked what type of WBCs is found in CTC-WBC clusters. We made use of a robotic micromanipulator to dissociate CTC-WBC clusters, enabling single-cell RNA sequencing analysis of cluster-associated WBCs and their comparison to reference WBCs from matched

donors (Fig. 1c) using reference component analysis (RCA) [71]. In patients, we found that 75% of CTC-associated WBCs relate to the myeloid lineage, while the remaining ones (25%) are similar to T cells (Fig. 1d and Extended Data Fig. 2a,b). Similarly, we found that 93% of CTC-associated WBCs from mouse models are also characterized by a myeloid cell-like expression profile (Extended Data Fig. 2c-e). To dissect the exact proportion of CTC-associated WBCs that are neutrophils, monocytes or macrophages, we labeled CTC-WBC clusters for Ly-6G, CD11b, F4/80, as well as with Wright-Giemsa staining to define nuclear morphology (Extended data Fig. 2f,g). We found that the vast majority (85.5-91.7%) of CTC-associated WBCs are positive for Ly-6G and display a nuclear morphology typical of neutrophils, while a minority (8.3-14.5%) are monocytes ($CD11b^+/F4/80^-/Ly\text{-}6G^-$) and no $F4/80^+$ macrophages are found (Fig. 1e,f and Extended Data Fig. 2h-j). Further, RNA sequencing analysis also revealed *ARG1*, *CXCL1*, *CXCL2*, *CXCL10*, *CCL2*, *CXCR2* and *VEGFA* expression in most CTC-associated neutrophils from both patients and mouse models (Extended Data Fig. 2k), indicating that CTC-associated neutrophils share gene expression features of pro-tumor N2-like cells [75].

We next asked whether the presence of CTC-neutrophil clusters in breast cancer patients could predict disease outcome. Strikingly, patients in whom at least one CTC-neutrophil cluster is detected in 7.5ml of peripheral blood are characterized by a significantly worse progression-free survival compared to patients with ≥ 5 CTCs per 7.5ml of peripheral blood (previously defined as a threshold for adverse outcome [76]) (Fig. 1g), as well as when compared to all patients with no CTC-neutrophil clusters, patients with at least one CTC per 7.5ml of blood, or patients in whom either single CTCs or CTC clusters are found (Extended Data Fig. 3a-c). Additionally, we individually micromanipulated equal numbers of CTCs from CTC-neutrophil clusters, CTC clusters and single CTCs, spontaneously generated from tumor-bearing mice, and intravenously injected 100 CTCs per mouse in tumor-free recipient mice from each of these categories. We found that mice injected with CTCs from CTC-neutrophil clusters develop overt metastasis much faster than those injected with CTCs alone, and accordingly, survive less (Fig. 1h and Extended Data Fig. 3d-h). Thus, CTC-neutrophil clusters represent the most efficient metastasis-forming cell subpopulation among breast CTCs, and their presence in the bloodstream of patients is associated with a poor prognosis.

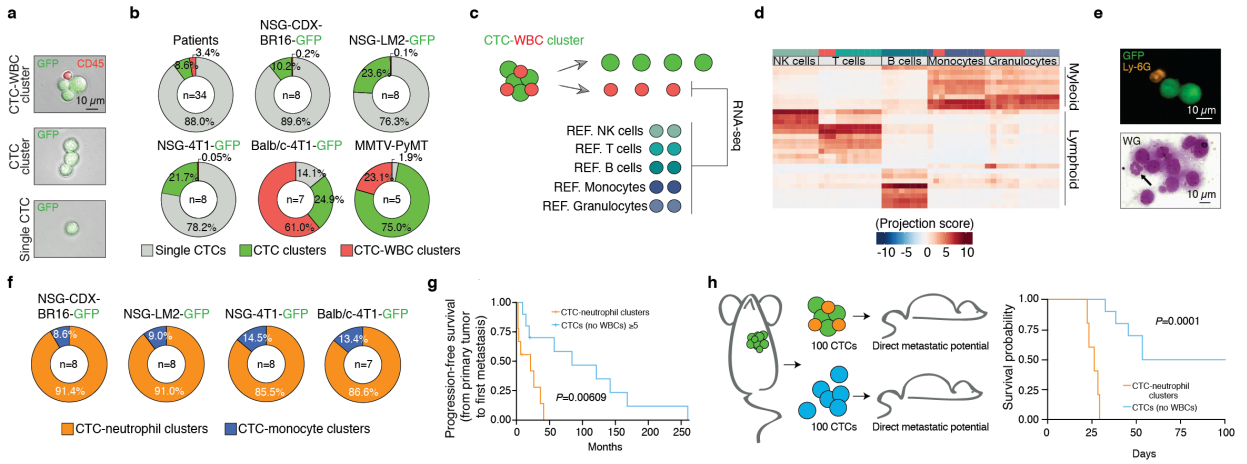


Figure 1. CTC-neutrophil clusters are highly-efficient metastatic precursors. a, Representative images of a CTC–WBC cluster, a CTC cluster and a single CTC from NSG-CDX-BR16-GFP mice. CTCs are expressing GFP (green) and CTC-associated WBCs are labelled with anti-CD45 antibodies (red) (n = 8). **b**, Pie charts displaying the mean percentage of single CTCs (grey), CTC clusters (green) and CTC–WBC clusters (red) in patients with breast cancer and in mouse models. The number of independent biological replicates (n) is shown for each model. **c**, Schematic of the experimental design. CTC–WBC clusters are dissociated into individual cells and processed for RNA sequencing (RNA-seq). The transcriptome of CTC-associated WBCs is then compared to reference (ref.) WBCs. **d**, Reference component analysis clustering of CTC-associated WBCs and reference WBCs from patients with breast cancer, displaying projection scores of cells (columns; n = 50) on the immune reference panel (rows). **e**, Representative images of CTC–WBC clusters stained for Ly-6G (neutrophils, gold) together with GFP (cancer cells, green) (top) or processed with the Wright–Giemsa (WG) assay to define nuclear morphology (bottom) (n = 8). **f**, Pie charts showing the mean percentage of CTC–neutrophil clusters and CTC–monocyte clusters in individual models. The number of independent biological replicates (n) is shown for each model. **g**, Kaplan–Meier plot showing progression-free survival of patients with breast cancer (n = 9 for patients with one or more CTC–neutrophil clusters and n = 10 for patients with five or more CTCs); P value by two-sided log-rank test is shown. **h**, Schematic of the experimental design. One hundred CTCs from CTC–neutrophil clusters or CTCs alone are injected in the tail vein of recipient mice to measure their metastatic potential (left). Kaplan–Meier plot showing overall survival of mice (right). n = 5 for CTC–neutrophil clusters and n = 10 for CTCs alone; P value by two-sided log-rank test is shown.

We next sought to determine the molecular consequences of the interaction between CTCs and neutrophils by dissociating CTC-neutrophil clusters and comparing the expression profile of CTCs from CTC-neutrophil clusters to that of CTCs alone (Fig. 2a). We first determined differential gene expression in the Balb/c-4T1-*GFP* model, where we could retrieve the highest

number of CTCs from CTC-neutrophil clusters (n=25). Compared to CTCs that were not associated to neutrophils, we found that CTCs from CTC-neutrophil clusters are characterized by differential expression of a set of 51 genes, of which 41 are upregulated and 10 are downregulated (Fig. 2b and Supplementary Tables 3,4). Pathway analysis with the upregulated genes revealed that CTCs from CTC-neutrophil clusters display a remarkable enrichment in positive regulators of cell cycle and DNA replication programs compared to CTCs alone (Fig. 2c), independently of the number of detected features or reads in each individual sample (Extended Data Fig. 4a). The same results were also obtained from CTCs of patients (Fig. 2c and Extended Data Fig. 4b). Immunofluorescence staining of CTCs confirmed higher levels of Ki67 expression in CTCs from CTC-neutrophil clusters (Fig. 2d,e), in line with the RNA sequencing results. In contrast, no significant differences were observed for genes involved in epithelial-to-mesenchymal transition, cancer stem cell markers, or platelet-related genes [77] (Extended Data Fig. 4c-h). Further, we asked whether neutrophil proximity would confer a proliferative advantage to cancer cells at the level of the primary tumor, disseminated tumor cells (DTCs) and overt metastasis. Surprisingly, we found that Ki67 expression does not increase in cancer cells that surround tumor-infiltrated neutrophils in the primary tumor or overt metastasis (Extended Data Fig. 5a-d). Yet, a higher Ki67 expression is retained in DTCs from CTC-neutrophil clusters (Extended Data Fig. 5e,f), i.e. when they are deprived of other stromal-derived signals that are typical of overt disease.

We then asked which cytokines are expressed by CTC-associated neutrophils and paralleled by simultaneous expression of matching cytokine receptor(s) in the corresponding cancer cells. We found that four cytokines (*TNFα*, *OSM*, *IL1β*, and *IL6*) are most frequently expressed by CTC-associated neutrophils of patients or patient-derived mouse models, and matched by the expression of their receptors by the corresponding cancer cells (Extended Data Fig. 6a). With a reverse approach, we also found that CTCs from CTC-neutrophil clusters most frequently express cytokines such as *CSF1*, *CSF3* (a.k.a. granulocyte colony-stimulating factor, G-CSF), *TGFβ3*, and *IL15*, possibly involved in neutrophil stimulation [78-81], while corresponding neutrophils express their receptors (Extended Data Fig. 6b). Consistently, we observed that a 24h *in vitro* treatment (coherent with neutrophil lifespan [82]) with IL6, IL1β or both was sufficient to confer proliferative advantage to 4T1 cells upon dissemination, leading to faster metastasis development and shorter overall survival of mice (Fig. 2f,g and Extended Data Fig. 6c-e). Further, CRISPR-mediated knockout of IL6 or IL1β receptors in cancer cells, namely

IL6ST and *IL1R1*, did not alter the frequency of spontaneously-generated CTC-neutrophil clusters but it suppressed their proliferative advantage (Extended Data Fig. 6f-h).

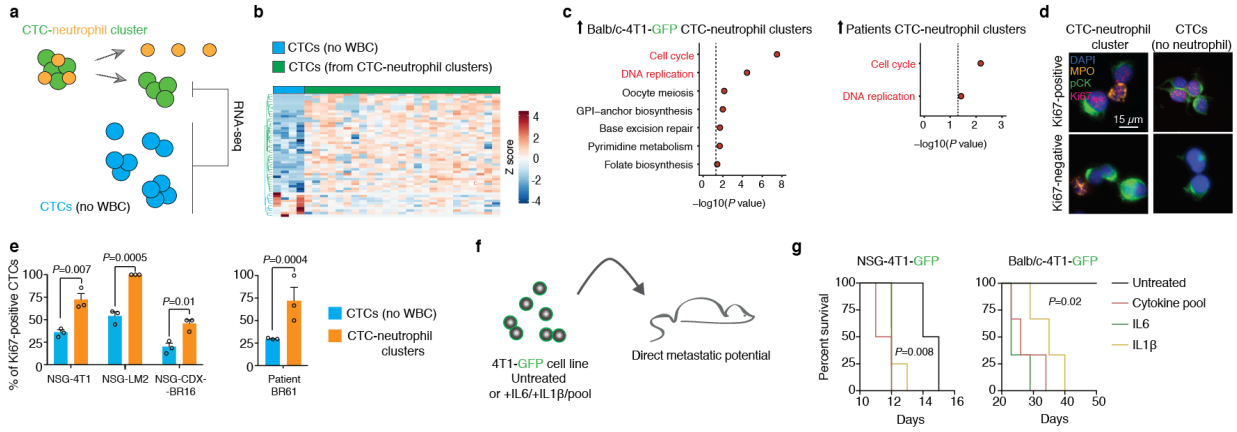


Figure 2. CTCs from CTC-neutrophil clusters are highly proliferative. a, Schematic of the experimental design. CTC–neutrophil clusters are dissociated and then processed for RNA sequencing. The transcriptome of CTCs from CTC–neutrophil clusters is compared to that of CTCs alone. **b**, Heat map of genes differentially expressed between CTCs from CTC–neutrophil clusters ($n = 25$) and CTCs alone ($n = 4$), isolated from BALB/c-4T1-*GFP* mice. The heat map displays gene-scaled (z-score) $\log_2$ counts per million mapped reads values after normalization, with columns representing samples ($n = 29$) and rows representing genes ($q < 0.05$ by edgeR likelihood ratio test). **c**, KEGG pathways over-represented ($P < 0.05$ by one-sided hypergeometric test) among upregulated genes in CTCs of CTC–neutrophil clusters from BALB/c-4T1-*GFP* mice (left) and test of selected pathways in patients ($P < 0.05$ by rotation gene set test (ROAST); right). **d**, Representative pictures of CTC–neutrophil clusters and CTCs alone from NSG-CDX-BR16-*GFP* mice, stained for pan-cytokeratin (pCK, green), myeloperoxidase (MPO, gold), Ki67 (purple) and DAPI (nuclei, blue) ($n = 3$). **e**, Plots showing the mean percentage of Ki67-positive CTCs. $n = 3$ for all; error bars, s.e.m.; P values by two-sided Student's t-test are shown. **f**, Schematic of the experimental design. 4T1-*GFP* cells are stimulated for 24 h with IL6, IL-1β or both (pool), then injected in recipient mice to assess their metastatic potential. **g**, Kaplan–Meier survival analysis of NSG (left) or BALB/c (right) mice injected with cytokine-treated 4T1-*GFP* cells. $n = 4$ for NSG-4T1-*GFP*, $n = 3$ for BALB/c-4T1-*GFP*, P values by two-sided log-rank test are shown.

Given recent findings highlighting that the presence of myeloid cells in the primary tumor site leads to accumulation of mutational events [83], we asked whether the mutational load of CTCs obtained from CTC-neutrophil clusters is different from that of CTCs alone in patients (Fig. 3a).

Interestingly, we found that the mutational burden is similar both between CTCs isolated from CTC-neutrophil clusters and CTC alone, as well as when comparing donors with or without CTC-neutrophil clusters (Extended Data Fig. 7a,b). Yet, we observed that the overall frequency of C>T mutations was increased in CTCs from CTC-neutrophil clusters compared to CTCs alone, as well as in donors with CTC-neutrophil clusters, independently of the nucleotide context (Fig. 3b and Extended Data Fig. 7c,d). While a general, age-related accumulation of C>T mutations has been reported [84], we did not observe any age difference between the two groups (Extended Data Fig. 7e). Next, considering only high-impact mutations, we asked whether specific genes are exclusively and recurrently mutated in donors with CTC-neutrophil clusters. This scenario would be consistent with a model whereby certain genetic alterations would influence the recruitment of immune cells to the primary tumor [85], and increase the likelihood of generating CTC-neutrophil clusters. We found that a number of genes are indeed carrying high-impact mutations exclusively in donors with CTC-neutrophil clusters (Extended Data Fig. 7f and Supplementary Table 5). We then engineered 4T1-*GFP* cells to individually express all the mutations found in two of the most frequently mutated genes, i.e. *MERTK* and *TLE1*, and injected them in the mammary gland of NSG mice. We observed that the introduction of *TLE1* mutations 1787G>A or 1509G>C leads to a higher infiltration of neutrophils in the primary tumor and a higher proportion of spontaneously-generated CTC-neutrophil clusters (33-41-fold increase), without affecting primary tumor size (Fig. 3c and Extended Data Fig. 7g-k). These results are in line with recent observations involving *TLE1* function in regulating myeloid cells infiltration into normal and neoplastic tissues[86]. Conversely, co-culture of cancer cells with neutrophils did not result in the acquisition of mutations within the same hotspots (Extended Data Fig. 8a,b). Together, our data reveals that while the overall mutational load remains unchanged, donors with CTC-neutrophil clusters feature a higher proportion of C>T substitutions and the presence of high-impact recurrent mutations in genes that promote neutrophil recruitment, such as *TLE1*.

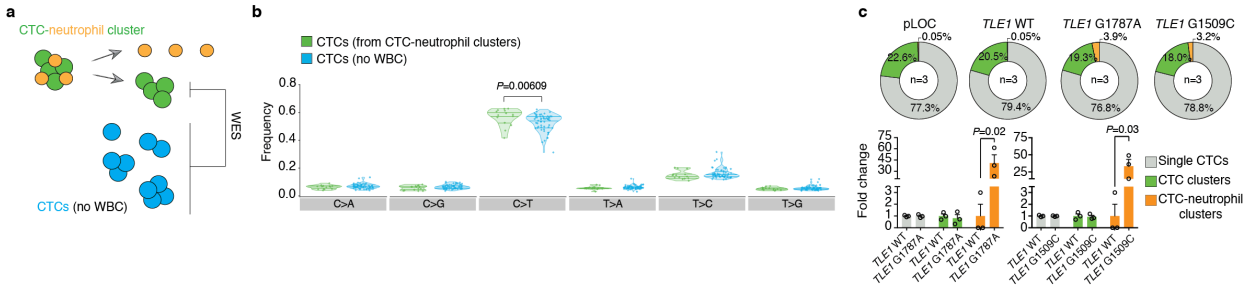


Figure 3. Whole-exome sequencing highlights recurrent mutational events in CTCs from CTC-neutrophil clusters. a, Schematic of the experimental design. CTC–neutrophil clusters are dissociated and then processed for whole-exome sequencing (WES). CTCs that were associated to neutrophils are compared to CTCs alone. **b**, Mutation distribution in CTCs from CTC–neutrophil clusters ($n = 14$) compared to CTCs alone ($n = 56$). Lines within the violin plots show the 25th, 50th and 75th percentile, respectively, and dots represent individual CTCs. *P* value by two-sided Wilcoxon sign-ranked test is shown. **c**, Pie charts displaying the mean percentage of single CTCs (grey), CTC clusters (green) and CTC–neutrophil clusters (gold) in mice injected with 4T1-pLOC, 4T1-*TLE1* WT, 4T1-*TLE1* G1787A or 4T1-*TLE1* G1509C cells (top). The number of independent biological replicates (*n*) is shown for each condition. The plots show the mean fold change of CTC ratios (bottom). $n = 3$; error bars, s.e.m.; *P* values by two-sided Student's *t*-test are shown.

We next tested whether the depletion or augmentation of the total neutrophil population would affect the ratio of spontaneously-generated CTC-neutrophil clusters. We depleted neutrophils by *in vivo* treatment with neutralizing antibodies against Ly-6G (αLy-6G) or conversely, we stably overexpressed G-CSF to stimulate the production and recruitment of neutrophils to the tumor site (Extended data Fig. 9a). As expected, while treatment with αLy-6G reduced neutrophil infiltration to the primary tumor site, G-CSF augmented it without altering primary tumor size (Extended Data Fig. 9b,c). Yet, αLy-6G-treated mice completely lack CTC-neutrophil clusters from the circulation and display a delayed CTC shedding rate compared to control mice, while overexpression of G-CSF leads to an earlier CTC release and substantially increases the proportion of CTC-neutrophil clusters (>88-fold) (Fig. 4a and Extended Data Fig. 9d-f). Consequently, neutrophil depletion or augmentation results in a delayed or accelerated metastasis development, respectively, mirrored by a shorter or longer overall survival of treated mice (Extended Data Fig. 9g,h). In contrast, neutrophil depletion with αLy-6G is not effective when cancer cells are administered directly through intravenous injection of pretreated mice (Extended Data Fig. 9i-l). Of note, in our cohort, G-CSF treatment of breast cancer patients

occurred more often in those patients that were positive for CTCs, including CTC-WBC clusters (Extended Data Fig. 9m). Thus, overall neutrophil abundance impacts the likelihood that a tumor has to spontaneously shed CTC-neutrophil clusters.

We next sought to identify actionable vulnerabilities of CTC-neutrophil clusters without targeting the entire neutrophil population. To this end, we investigated cell-cell junction pairs expressed by CTC-neutrophil clusters and possibly mediating their heterotypic cell binding (Extended Data Fig. 10a,b). We engineered a CRISPR/Cas9-based loss-of-function minipool screen *in vivo*, whereby a pool of cells carrying individual knockouts for *F11R*, *ICAM1*, *ITGB2* and *VCAM1* are injected in the mammary gland of recipient mice, followed by CTC targeted barcode sequencing to reveal selective sgRNA dropouts, highlighting genes whose knockout does not allow CTC-neutrophil clusters formation (Fig. 4b). Importantly, we observed no differences in primary tumor growth and no selective sgRNA dropouts in primary tumor cells (Extended Data Fig. 10c,d), suggesting that knockout of *F11R*, *ICAM1*, *ITGB2* or *VCAM1* does not affect proliferation in the primary tumor. Yet, we found that 4/4 *VCAM1* sgRNAs selectively dropped out in CTCs from CTC-neutrophil clusters, while they were still present in CTCs alone (Fig. 4c), highlighting a possible VCAM1 requirement for CTC-neutrophil clusters formation. We further validated this finding using individual sgRNAs (Fig. 4d and Extended Data Fig. 10e). Thus, VCAM1 functionally mediates the interaction between CTCs and neutrophils, and its inhibition does not allow the formation of CTC-neutrophil clusters.

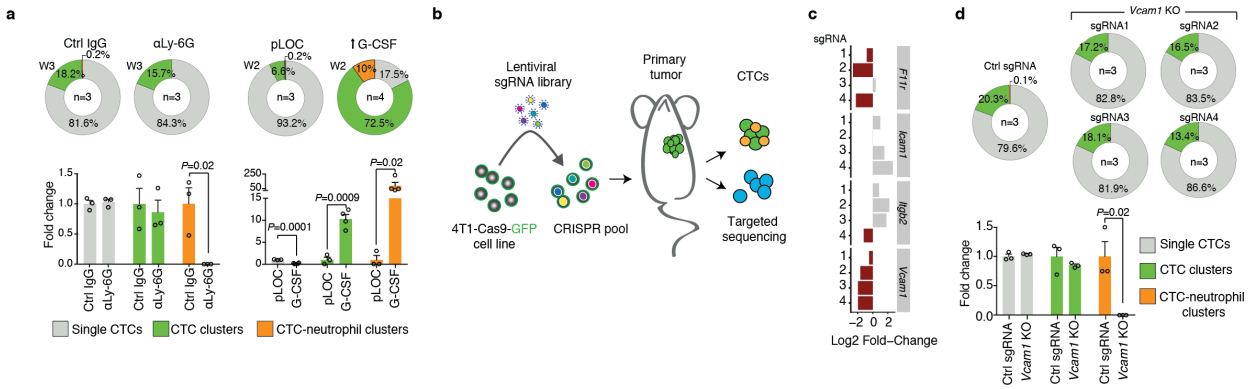


Figure 4. Identification of vulnerabilities of CTC-neutrophil clusters. a, Pie charts displaying the mean percentage of single CTCs (grey), CTC clusters (green) and CTC–neutrophil clusters (gold) in NSG-4T1-*GFP* mice treated with anti-Ly-6G antibodies or G-CSF overexpression (top). W, weeks upon tumour development. The number of independent biological replicates (*n*) is shown for each condition. The plots show the mean fold change

of CTC ratios from NSG-4T1-*GFP* mice treated with anti-Ly-6G antibodies or G-CSF overexpression (bottom). Error bars, s.e.m.; *P* values by two-sided Student's *t*-test are shown. **b**, Schematic of the experimental design. 4T1-Cas9-*GFP* cells are transduced with a pool of sgRNAs targeting cell–cell junctions and injected in NSG mice. Upon tumour development, spontaneously generated CTC–neutrophil clusters, CTC clusters and single CTCs are sequenced to identify sgRNA dropouts. **c**, Bar plot showing the fold change ($\log_2$) of individual sgRNAs (numbered 1 to 4) found in CTCs from CTC–neutrophil clusters versus CTCs alone. sgRNAs with a representation that was reduced in CTCs from CTC–neutrophil clusters are shown in dark red ($n = 3$). **d**, Pie charts displaying the mean percentage of single CTCs (grey), CTC clusters (green) and CTC–neutrophil clusters (gold) in NSG mice carrying 4T1-Cas9-*GFP* tumours expressing a control sgRNA or individual sgRNAs targeting *Vcam1* (*Vcam1*-KO) (top). The number of independent biological replicates (n) is shown for each condition. The plot shows the mean fold change of CTC ratios from mice carrying 4T1-Cas9-*GFP* tumours expressing a control sgRNA or individual sgRNAs targeting *Vcam1* (bottom). Error bars, s.e.m., *P* value by two-sided Student's *t*-test is shown.

Altogether, our data provide new insights into the processes that define the interaction between cancer cells and immune cells during blood-borne dissemination (Fig. 5) We propose a model whereby neutrophils directly interact with CTCs to support cell cycle progression in circulation and to accelerate metastasis seeding. This mechanism of metastatic spread and the possibility that CTC-neutrophil clusters may be targeted therapeutically provides an opportunity to reduce the spread of breast cancer.

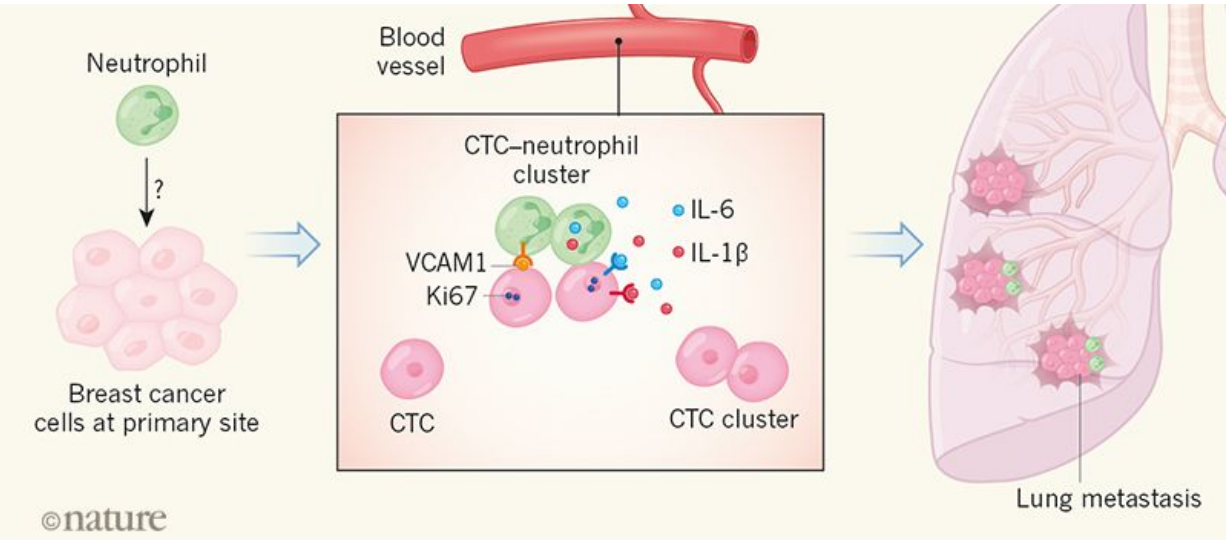
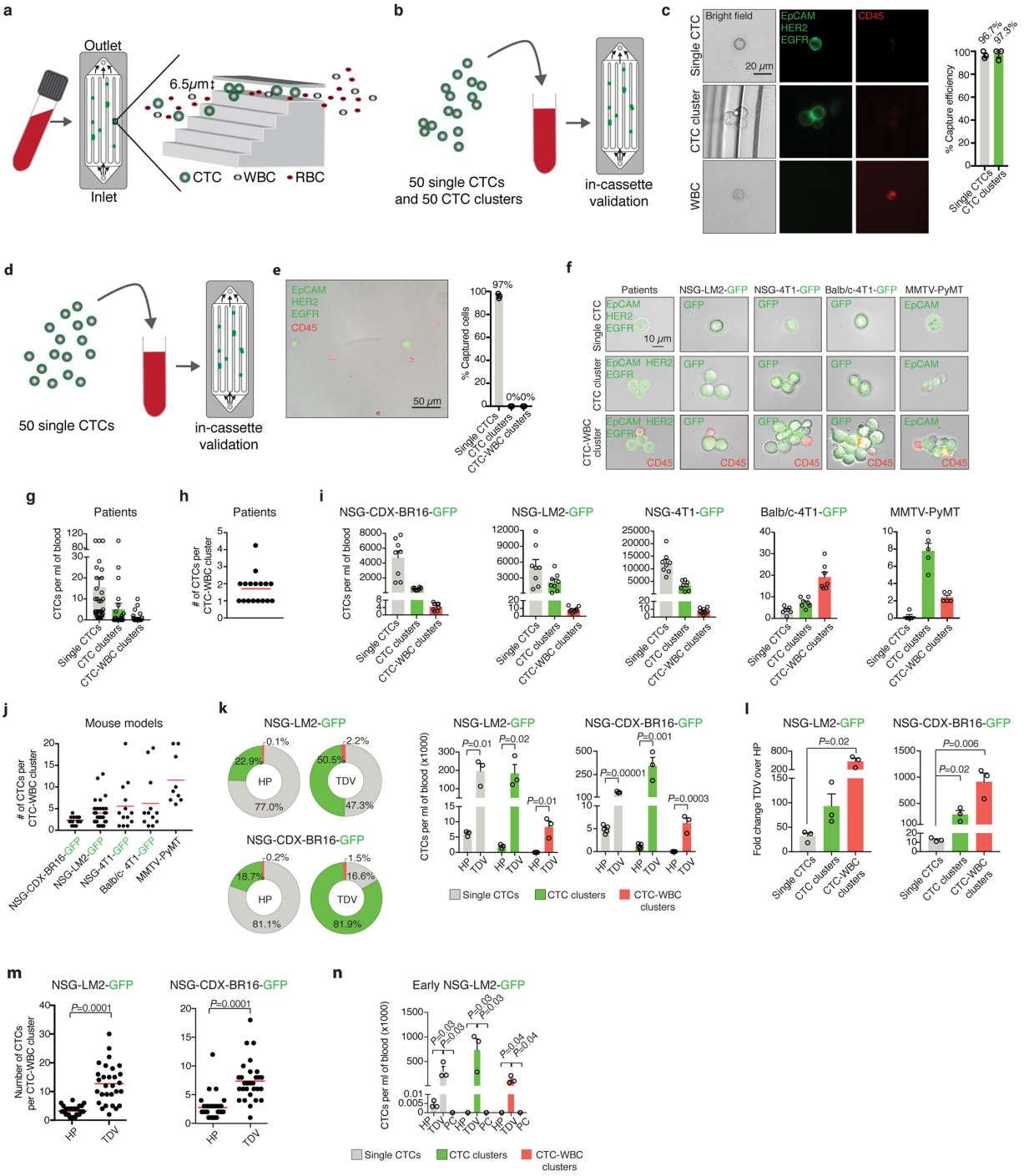
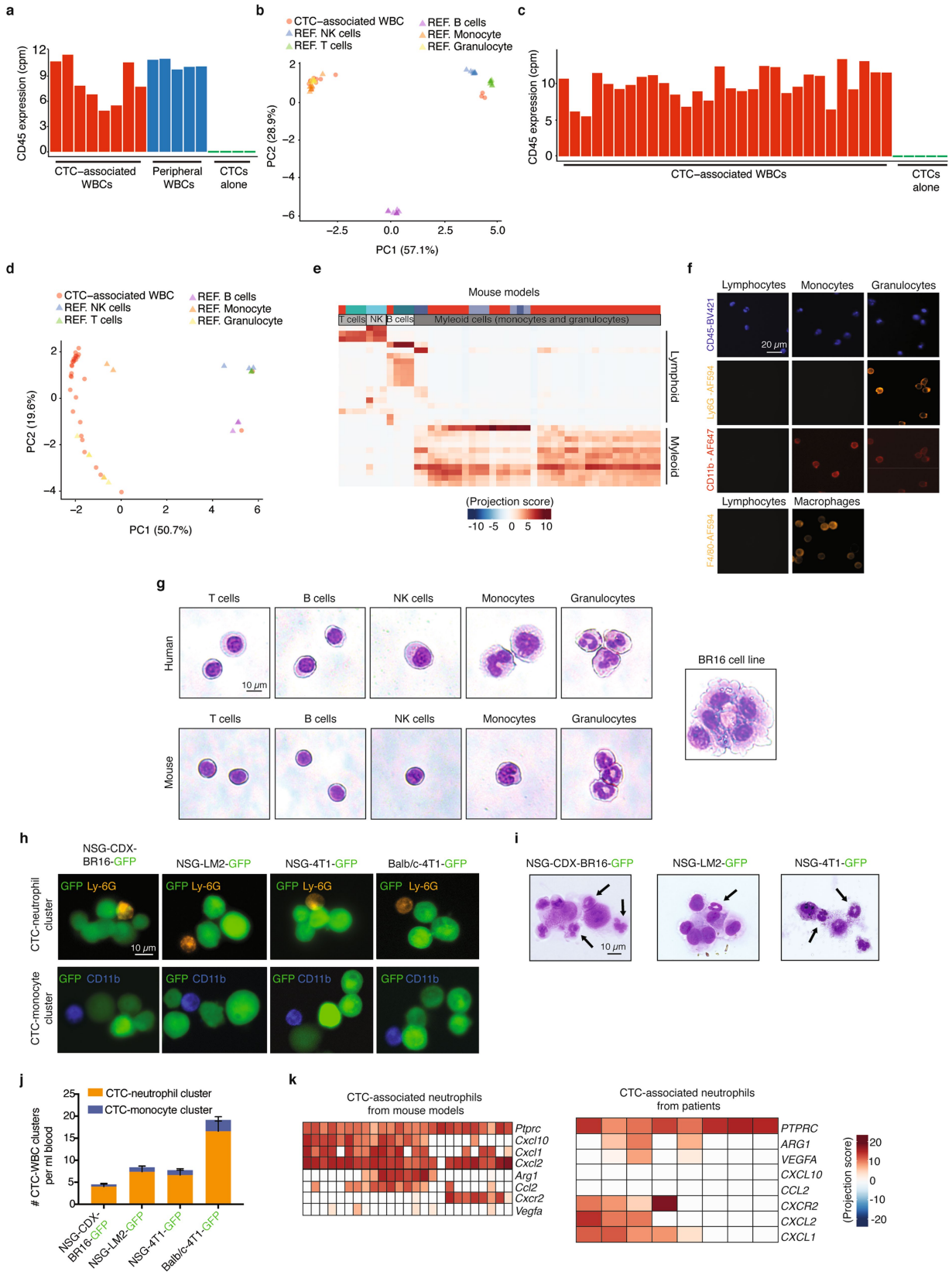


Figure 5. Graphic representation of key findings. Image from 'Sticking together helps cancer to spread' by Egeblad and de Visser, Nature 2019 [87].



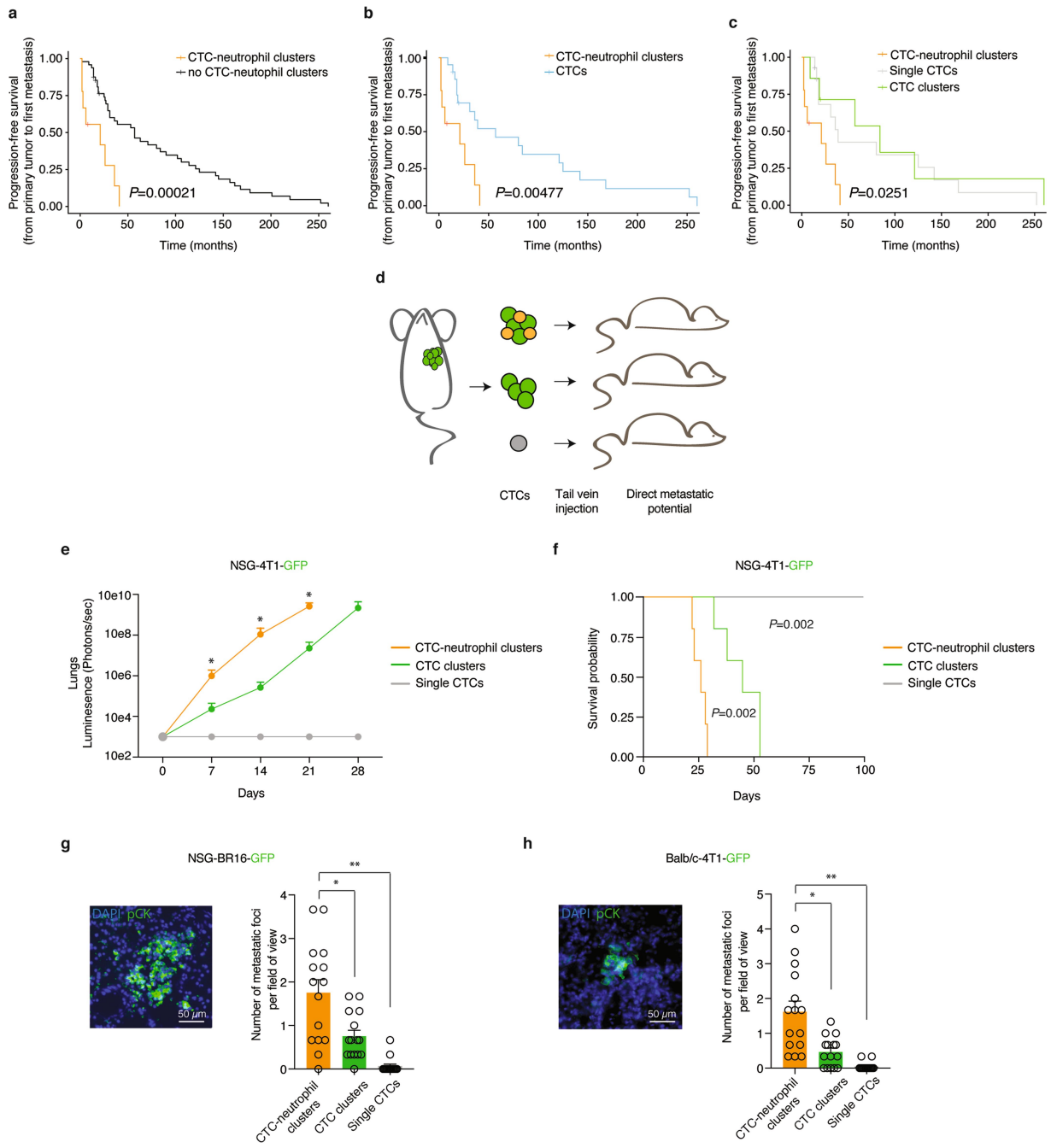
Extended Data Figure 1. CTC capture in breast cancer patients and mouse models. a, Schematic of the CTC capture strategy with the Parsortix device. **b,** Schematic of the experimental design. 50 single CTCs and 50 CTC clusters are spiked into blood to assess capture rate. **c,** Representative pictures of CTCs and WBCs captured on the Parsortix device and stained for EpCAM, HER2, EGFR (*green*) and CD45 (*red*) (*left*); *n*=3. The plot shows the mean CTC capture efficiency (*right*); *n*=3; error bars represent S.E.M. **d,** Schematic of the experimental design. 50

single CTCs are spiked into blood to evaluate artificial CTC aggregation rate during processing. **e,** Representative picture of captured CTCs (*left*); *n*=3. The plot shows the mean percent of captured single CTCs, CTC clusters and CTC-WBC clusters (*right*); *n*=3; error bars represent S.E.M. **f,** Representative pictures of CTCs from patients and mouse models. *n*=34 for patients; *n*=8 for NSG-LM2-*GFP* and NSG-4T1-*GFP*; *n*=7 for Balb/c-4T1-*GFP*; *n*=5 for MMTV-PyMT. **g,** The plot shows mean CTC counts in patients. *n*=34; error bars represent S.E.M. **h,** The plot shows the number of CTCs in each patient-derived CTC-WBC cluster. The *red* line represents the mean. **i,** The plots show mean CTC counts in mouse models. *n*=8 for NSG-CDX-BR16-*GFP*, NSG-LM2-*GFP* and NSG-4T1-*GFP*; *n*=7 for Balb/c-4T1-*GFP*; *n*=5 for MMTV-PyMT. Error bars represent S.E.M. **j,** The plot shows the number of CTCs in each mouse model-derived CTC-WBC cluster. The *red* line represents the mean. **k,** Pie charts displaying the mean percentage of single CTCs (*grey*), CTC clusters (*green*) and CTC-WBC clusters (*red*) in mice upon blood draw *via* heart puncture (HP) or tumor-draining vessel (TDV) (*left*). The plots show the mean number of CTCs from the same experiment (*right*). Error bars represent S.E.M.; *n*=5 for NSG-CDX-BR16-*GFP*, *n*=3 for NSG-LM2-*GFP*; *P* values by two-sided Student's *t* test are shown. **l,** The plots show fold change of CTC counts, comparing HP *versus* TDV blood draw. Error bars represent S.E.M.; *n*=5 for NSG-CDX-BR16-*GFP*, *n*=3 for NSG-LM2-*GFP*; *P* values by two-sided Student's *t* test are shown. **m,** The plots show the number of CTCs in each mouse model-derived CTC-WBC cluster, isolated *via* HP or TDV. The *red* lines represent the mean. *P* values by two-sided Student's *t* test are shown. **n,** The plot shows the mean number of CTCs at day 10 after tumor inoculation, collected from HP, TDV or peripheral circulation (i.e. tail vein; PC). Error bars represent S.E.M.; *n*=3; *P* values by two-sided Student's *t* test are shown.



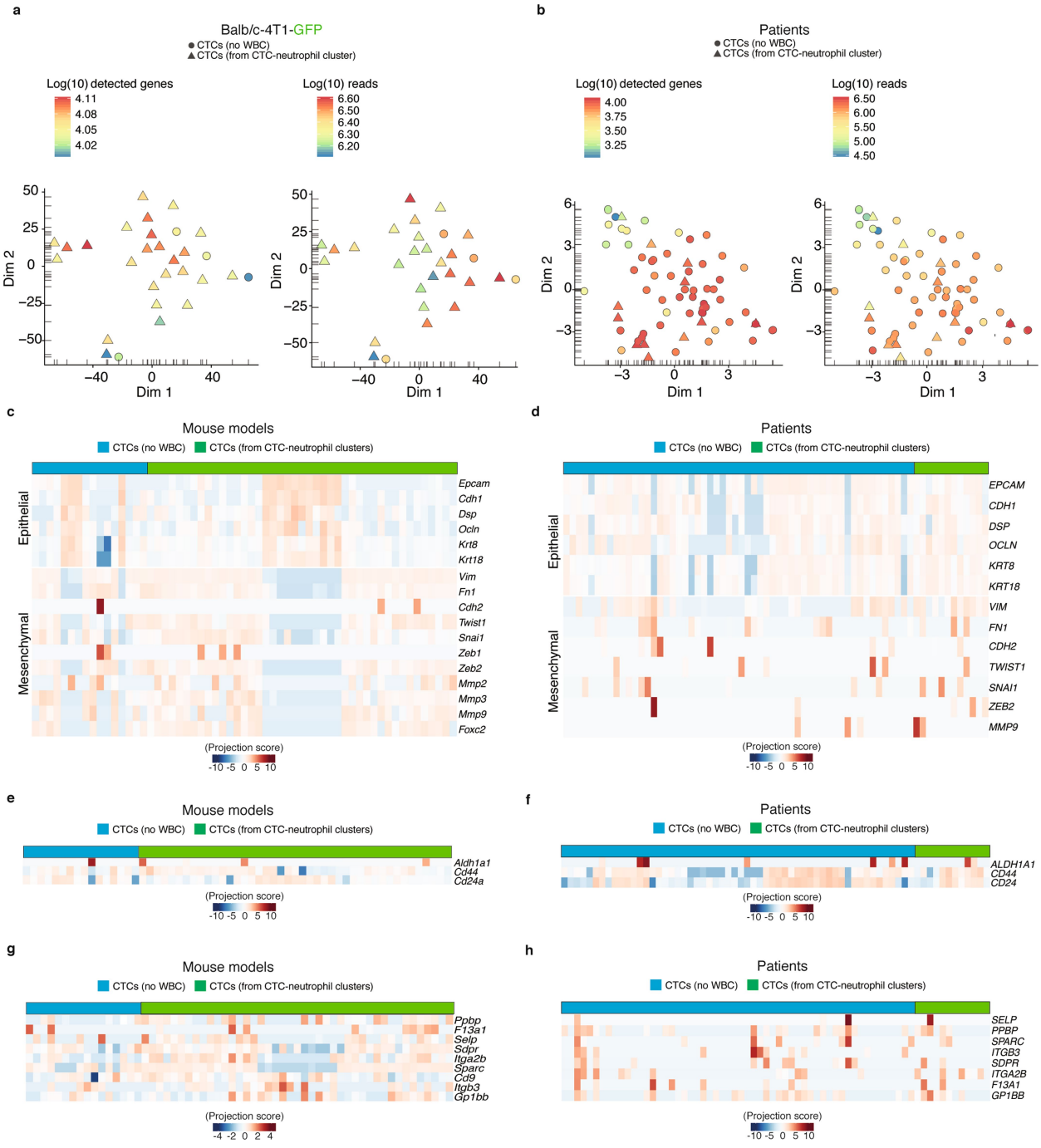
Extended Data Figure 2. Characterization of CTC-associated WBCs. a, Bar plot showing the expression levels of white blood cell (WBC) marker CD45 in patient samples, including CTC-associated WBCs (*red*), free-floating peripheral WBCs (*blue*) and CTCs alone (*green*). **b,** Principal component analysis (PCA) of CTC-associated WBCs

of patients and five reference WBC populations. *n*=50. **c,** Bar plot showing the expression levels of CD45 in mouse samples, including CTC-associated WBCs (*red*) and CTCs alone (*green*). **d,** PCA of CTC-associated WBCs from all mouse models and five reference WBC populations. *n*=47. **e,** Reference component analysis clustering of CTC-associated WBCs (*red*) and reference WBCs from mouse models, displaying projection scores of cells (*columns; n=47*) on the immune reference panel (*rows*). **f**, Representative immunofluorescence images of mouse immune cells stained for CD45, Ly-6G and CD11b. Mouse lymphocytes ($CD45^+/Ly\text{-}6G^-/CD11b^-$), mouse monocytes ($CD45^+/Ly\text{-}6G^-/CD11b^+$), and mouse granulocytes ($CD45^+/Ly\text{-}6G^+/CD11b^{low}$) are shown (*top*). Representative immunofluorescence images of mouse lymphocytes and macrophages (peritoneum-derived) stained for F4/80 are shown (*bottom*). Mouse macrophages display a $F4/80^+$ phenotype, while lymphocytes display a $F4/80^-$ phenotype. *n*=3 for all. **g,** Representative images of human and mouse T cells, B cells, NK cells, monocytes and granulocytes stained with the Wright-Giemsa protocol to highlight nuclear morphology (*left*). Wright-Giemsa staining of the human CTC-derived cell line BR16 is also shown (*right*). *n*=3 for all. **h,** Representative immunofluorescence images of CTC-neutrophil clusters and CTC-monocyte clusters isolated from mouse models and stained for Ly-6G (*gold*) and CD11b (*blue*). CTCs stably express *GFP* (*green*). *n*=8 for NSG-CDX-BR16-*GFP*, NSG-LM2-*GFP* and NSG-4T1-*GFP*; *n*=7 for Balb/c-4T1-*GFP*. **i,** Representative images of CTC-neutrophil clusters stained with the Wright-Giemsa protocol to highlight nuclear morphology. *n*=8 for NSG-CDX-BR16-*GFP*, NSG-LM2-*GFP* and NSG-4T1-*GFP*. **j,** Bar graph showing the mean number of CTC-neutrophil clusters and CTC-monocyte clusters in mouse models. Error bars represent S.E.M.; *n*=8 for NSG-CDX-BR16-*GFP*, NSG-LM2-*GFP*, NSG-4T1-*GFP* and *n*=7 for Balb/c-4T1-*GFP*. **k,** Heatmaps showing the projection scores of mouse-derived (*left*) and patient-derived (*right*) CTC-associated neutrophils (*columns*) on pro-tumoral (N2) neutrophil markers (*rows*).



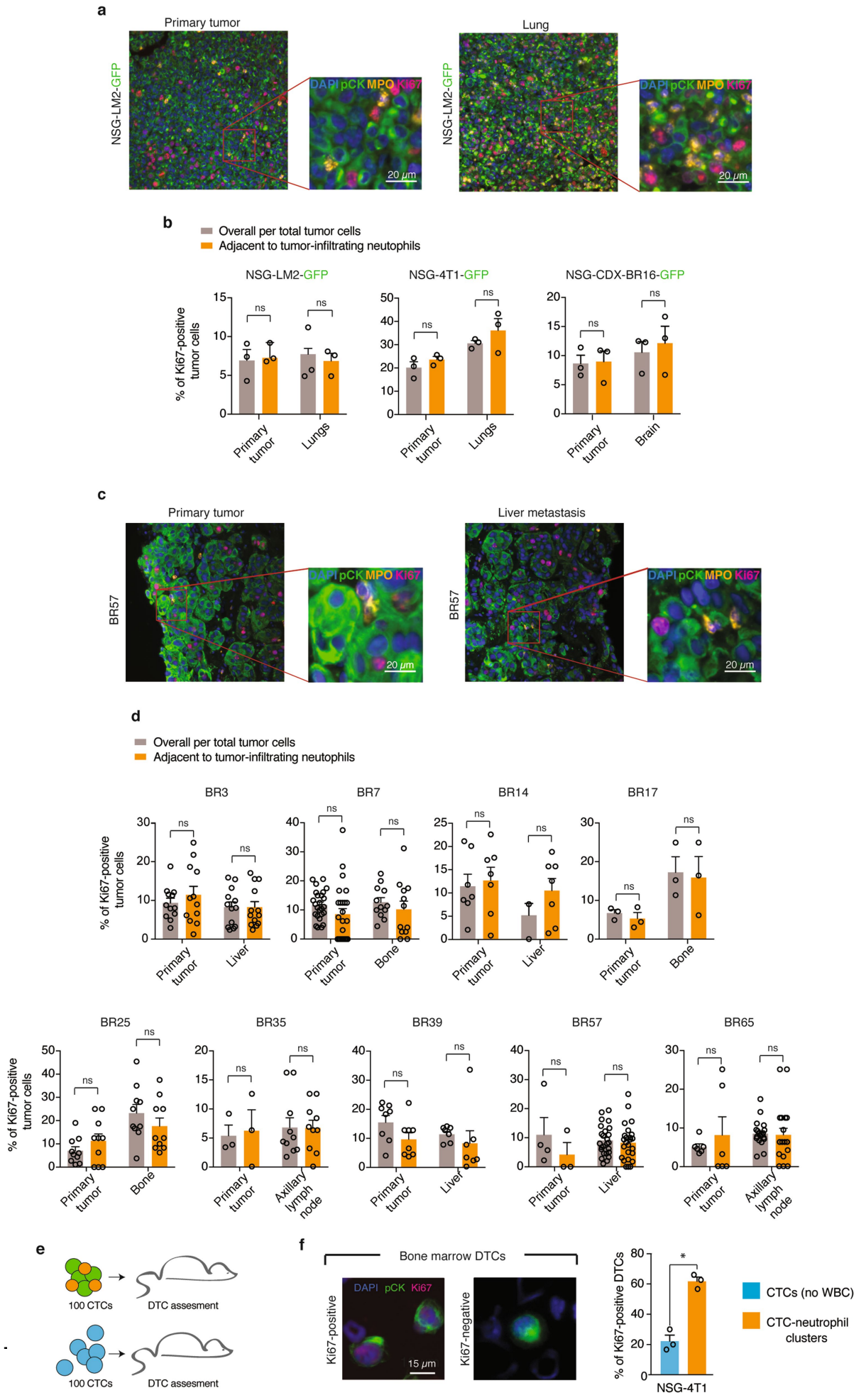
Extended Data Figure 3. Progression-free survival analysis in breast cancer patients and mouse models. **a,** Kaplan-Meier progression-free survival (PFS) analysis comparing patients with ≥1 CTC-neutrophil cluster per 7.5ml of peripheral blood (*n*=9) *versus* all patients with no CTC-neutrophil clusters (*n*=48); *P* value by two-sided Log-Rank test is shown. **b,** Kaplan-Meier PFS analysis comparing patients with ≥1 CTC-neutrophil cluster per 7.5ml of peripheral blood (*n*=9) *versus* patients with ≥1 CTC per 7.5ml of peripheral blood but no CTC-neutrophil clusters (*n*=21); *P* value by two-sided Log-Rank test is shown. **c,** Kaplan-Meier PFS analysis comparing patients with ≥1 CTC-neutrophil cluster per 7.5ml of peripheral blood (*n*=9), patients with ≥1 single CTC per 7.5ml of peripheral blood but without CTC-neutrophil clusters (*n*=14), and patients with ≥1 CTC cluster per 7.5ml of

peripheral blood but without CTC-neutrophil clusters (*n*=7).; *P* value by two-sided Log-Rank test is shown. Of note, these results are consistent with our previous observations whereby PFS differences in patients with single CTCs *versus* CTC-clusters were visible only when CTC clusters were present for multiple time points along disease progression. **d,** Schematic of the experimental design. 100 CTCs from CTC-neutrophil clusters, CTC clusters or single CTCs are injected in the tail vein of tumor-free recipient mice to measure their metastatic potential. **e,** The plot shows normalized bioluminescence signal from the lungs of injected mice. *n*=5 for all; Error bars represent S.E.M.; **P*< 0.05 by two-sided Student's *t* test. **f,** Kaplan-Meier plot showing overall survival of injected mice. *n*=5 for all; *P* values by two-sided Log-Rank test are shown. **g,** Representative picture of a metastatic lesion in NSG mice injected intravenously with either with CTC-neutrophil clusters, CTC clusters or single CTCs from NSG-BR16-*GFP* mice. Metastases are stained for pan-cytokeratin (pCK; *green*) and DAPI (nuclei; *blue*) (*left*); *n*=3. The plot shows the mean number of metastatic foci per field of view (*right*). *n*=3; Error bars represent S.E.M; *P* values by two-sided Student's *t* test are shown. **h,** Representative picture of a metastatic lesion in the lungs of Balb/c mice injected intravenously with either with CTC-neutrophil clusters, CTC clusters or single CTCs from Balb/c-4T1-*GFP* mice. Metastases are stained for pan-cytokeratin (pCK; *green*) and DAPI (nuclei; *blue*) (*left*); *n*=3. The plot shows the mean number of metastatic foci per field of view (*right*). *n*=3; Error bars represent S.E.M; *P* values two-sided Student's *t* test are shown.

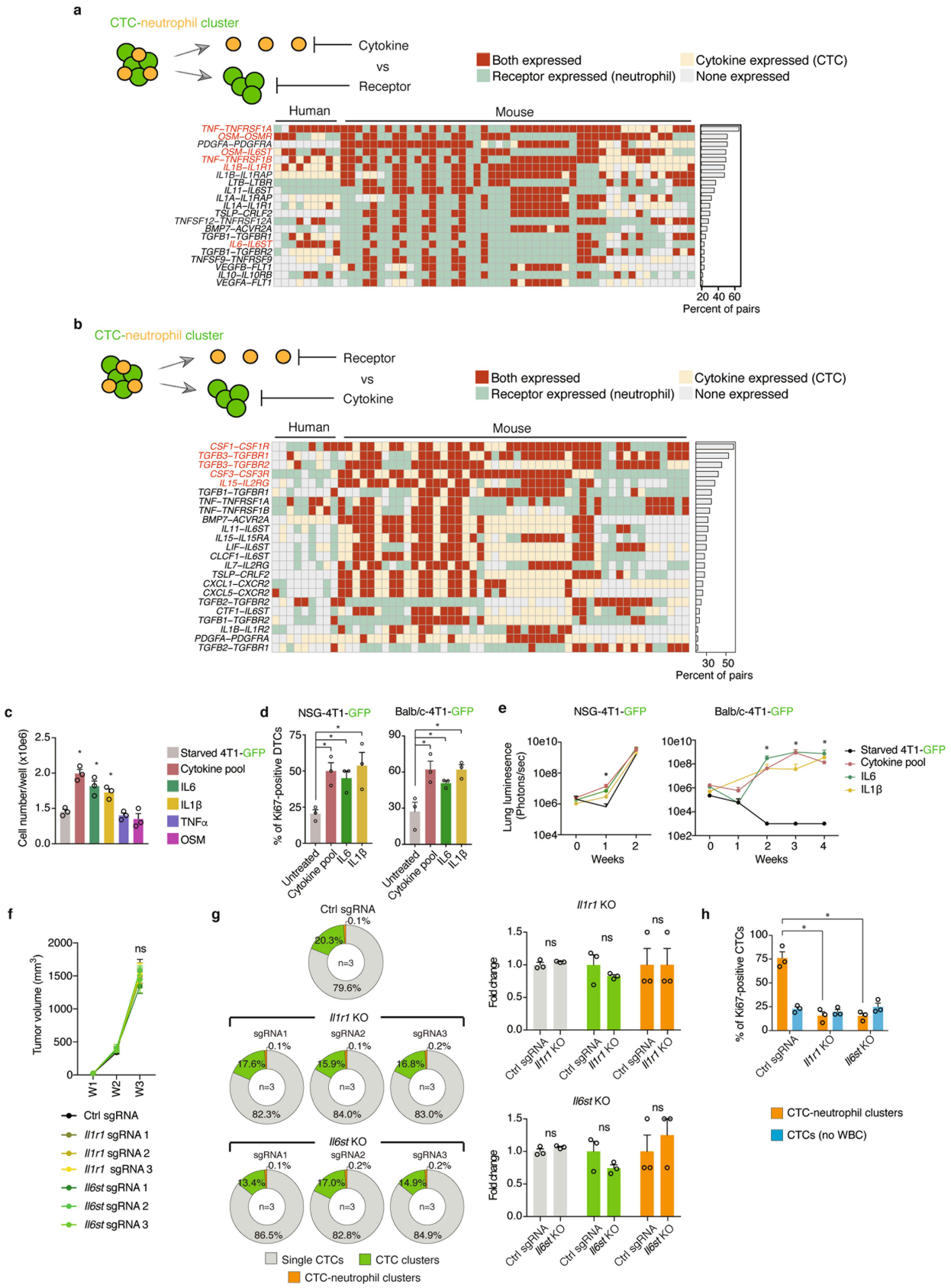


Extended Data Figure 4. Gene expression analysis of single-cell RNA sequencing data. a, t-distributed stochastic neighbor embedding (t-SNE) analysis of CTCs from CTC-neutrophil clusters and CTCs alone using the 500 most variable genes. t-SNE plots for Balb/c-4T1-*GFP* samples are colored by number of detected genes (*left*) and number of reads per sample (*right*). *n*=29. **b,** t-SNE plots for patient samples colored by number of detected genes (*left*) and number of reads per sample (*right*). *n*=68. **c,** Heatmap showing the projection scores of mouse model-derived CTCs from CTC-neutrophil clusters and CTCs alone in relation to epithelial and mesenchymal genes. *n*=59. **d,** Heatmap showing the projection scores of patient-derived CTCs from CTC-neutrophil clusters and

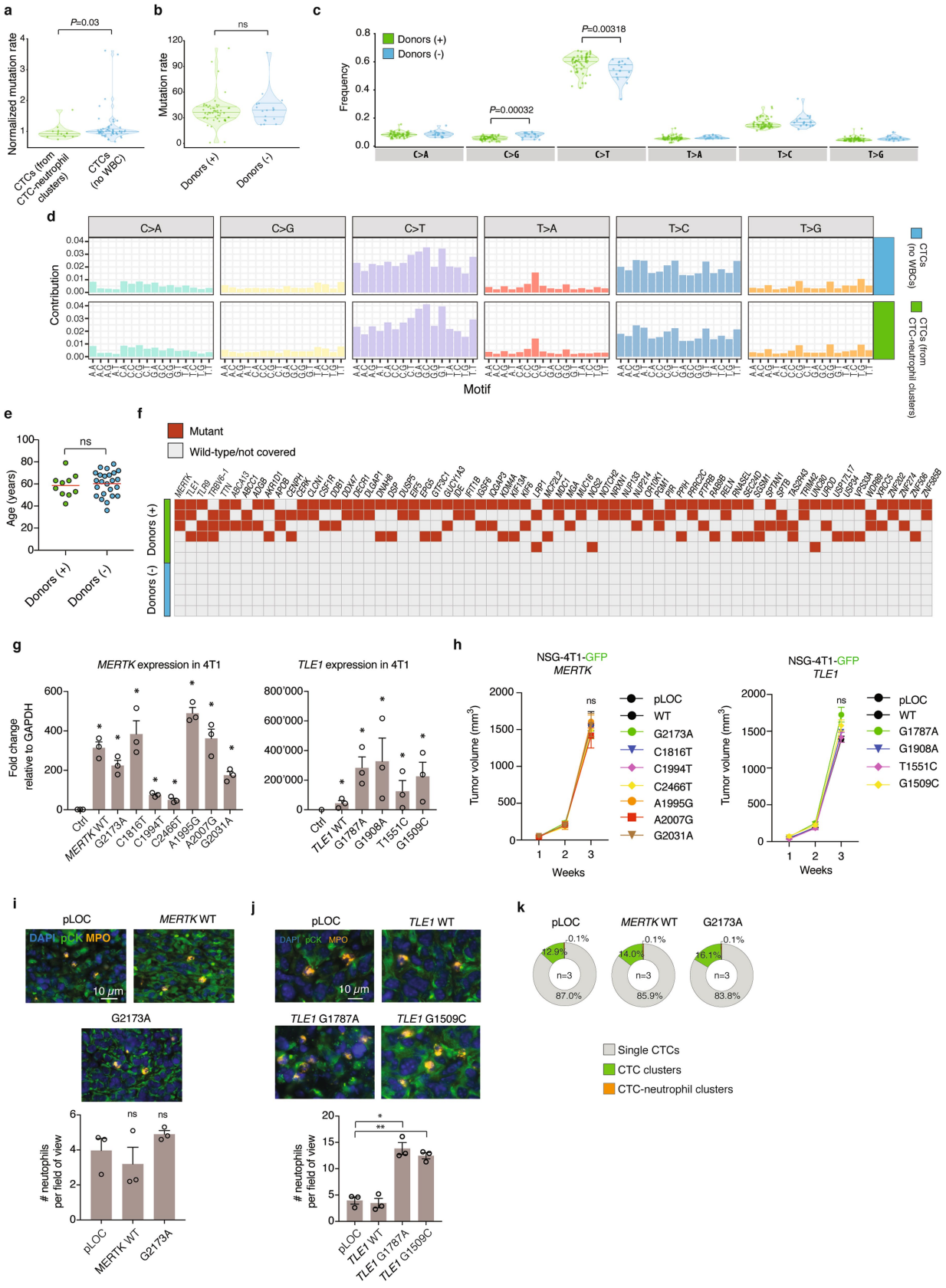
CTCs alone in relation to epithelial and mesenchymal genes. *n*=68. **e,** Heatmap showing the projection scores of mouse model-derived CTCs from CTC-neutrophil clusters and CTCs alone in relation to cancer stem cell genes. *n*=59. **f,** Heatmap showing the projection scores of patient-derived CTCs from CTC-neutrophil clusters and CTCs alone in relation to cancer stem cell genes. *n*=68. **g,** Heatmap showing the projection scores of mouse model-derived CTCs from CTC-neutrophil clusters and CTCs alone in relation to platelet genes. *n*=59. **h,** Heatmap showing the projection scores of patient-derived CTCs from CTC-neutrophil clusters and CTCs alone in relation to platelet genes. *n*=68.


a
Primary tumor
Lung
NSG-LM2-GFP
DAPI pCK MPO Ki67
20 μm
b
Overall per total tumor cells
Adjacent to tumor-infiltrating neutophils
NSG-LM2-GFP
NSG-4T1-GFP
NSG-CDX-BR16-GFP
% of Ki67-positive tumor cells
ns
Primary tumor
Lungs
Brain
c
Primary tumor
Liver metastasis
BR57
d
BR3
BR7
BR14
BR17
BR25
BR35
BR39
BR57
BR65
Liver
Bone
Axillary lymph node
e
100 CTCs
DTC assesment
f
Bone marrow DTCs
Ki67-positive
Ki67-negative
15 μm
% of Ki67-positive DTCs
CTCs (no WBC)
CTC-neutrophil clusters
NSG-4T1

Extended Data Figure 5. Proliferation of tumor cells adjacent to neutrophils in primary and metastatic tissues. a, Representative immunofluorescence images of NSG-LM2-*GFP* primary tumor and matched lung metastasis, stained for pan cytokeratin (pCK; *green*), MPO (*gold*), Ki67 (*purple*) and DAPI (nuclei; *blue*). *n*=3. **b,** The plots show the mean percent of Ki67-positive cancer cells in the primary tumor and metastatic sites (lung or brain) of mouse models, both overall and when considering only those cells that are adjacent to neutrophils. *n*=3 for all; error bars represent S.E.M; ns= not significant by two-sided Student's *t* test. **c,** Representative immunofluorescence images of BR57 primary tumor and matched liver metastasis, stained for pan cytokeratin (pCK; *green*), MPO (*gold*), Ki67 (*purple*) and DAPI (nuclei; *blue*). *n*=3. **d,** The plots show the mean percent of Ki67-positive cancer cells, both overall and when considering only those cells that are adjacent to neutrophils, in matched primary and metastatic sites of nine breast cancer patients. Error bars represent S.E.M.; ns= not significant by two-sided Student's *t* test. **e,** Schematic of the experimental design. 100 CTCs from CTC-neutrophil clusters or CTC alone are injected in the tail vein of recipient mice to measure disseminated tumor cells (DTC) proliferation. **f,** Representative pictures of DTCs stained for pan-cytokeratin (pCK; *green*), Ki67 (*purple*) and DAPI (nuclei; *blue*) (*left*). The plot shows the mean percent of Ki67-positive DTCs (*right*). *n*=3. Error bars represent S.E.M; **P*=0.001 by two-sided Student's *t* test.

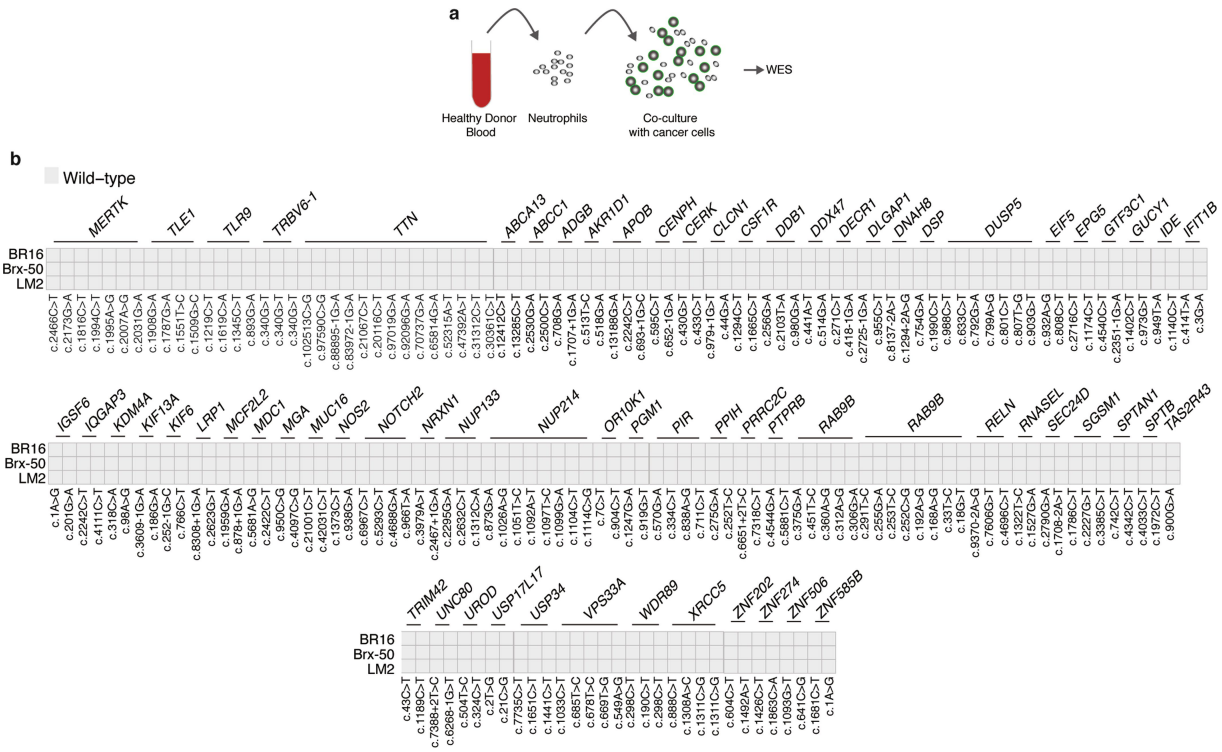

a
CTC-neutrophil cluster
Cytokine
vs
Receptor
Both expressed
Receptor expressed (neutrophil)
Cytokine expressed (CTC)
None expressed
Human
Mouse
TNF–TNFRSF1A
OSM–OSMR
PDGFA–PDGFRA
OSM–IL6ST
TNF–TNFRSF1B
IL1B–IL1R1
IL1B–IL1RAP
LTB–LTBR
IL11–IL6ST
IL1A–IL1RAP
IL1A–IL1R1
TSLP–CRLF2
TNFSF12–TNFRSF12A
BMP7–ACVR2A
TGFB1–TGFBR1
IL6–IL6ST
TGFB1–TGFBR2
TNFSF9–TNFRSF9
VEGFB–FLT1
IL10–IL10RB
VEGFA–FLT1
20 40 60
Percent of pairs
b
CTC-neutrophil cluster
Receptor
vs
Cytokine
Both expressed
Receptor expressed (neutrophil)
Cytokine expressed (CTC)
None expressed
Human
Mouse
CSF1–CSF1R
TGFB3–TGFBR1
TGFB3–TGFBR2
CSF3–CSF3R
IL15–IL2RG
TGFB1–TGFBR1
TNF–TNFRSF1A
TNF–TNFRSF1B
BMP7–ACVR2A
IL11–IL6ST
IL15–IL15RA
LIF–IL6ST
CLCF1–IL6ST
IL7–IL2RG
TSLP–CRLF2
CXCL1–CXCR2
CXCL5–CXCR2
TGFB2–TGFBR2
CTF1–IL6ST
TGFB1–TGFBR2
IL1B–IL1R2
PDGFA–PDGFRA
TGFB2–TGFBR1
30 50
Percent of pairs
c
Cell number/well (x10e6)
Starved 4T1-GFP
Cytokine pool
IL6
IL1β
TNFα
OSM
d
NSG-4T1-GFP
Balb/c-4T1-GFP
% of Ki67-positive DTCs
Untreated
Cytokine pool
IL6
IL1β
e
NSG-4T1-GFP
Balb/c-4T1-GFP
Lung luminescence (Photons/sec)
Weeks
Starved 4T1-GFP
Cytokine pool
IL6
IL1β
f
Tumor volume (mm3)
ns
W1
W2
W3
Ctrl sgRNA
Il1r1 sgRNA 1
Il1r1 sgRNA 2
Il1r1 sgRNA 3
Il6st sgRNA 1
Il6st sgRNA 2
Il6st sgRNA 3
g
Ctrl sgRNA
20.3%
0.1%
n=3
79.6%
Il1r1 KO
sgRNA1
sgRNA2
sgRNA3
17.6%
15.9%
16.8%
0.1%
0.1%
0.2%
82.3%
84.0%
83.0%
Il6st KO
13.4%
17.0%
14.9%
0.2%
0.2%
86.5%
82.8%
84.9%
Single CTCs
CTC clusters
CTC-neutrophil clusters
Il1r1 KO
Il6st KO
Fold change
ns
Ctrl sgRNA
Il1r1 KO
Il6st KO
h
% of Ki67-positive CTCs
CTC-neutrophil clusters
CTCs (no WBC)

Extended Data Figure 6. Characterization of cytokine-mediated crosstalk within CTC-neutrophil clusters. a, Schematic of the experimental design (*top*). The heatmap shows the transcriptional landscape of cytokines and corresponding receptors expressed in at least 20% of CTC-neutrophil clusters (*bottom*). The cytokine-receptor pairs that are most frequently expressed in human cells are shown in *red*. **b,** Schematic of the experimental design (*top*). The heatmap shows the transcriptional landscape of cytokine receptors and corresponding cytokines expressed in at least 20% of CTC-neutrophil clusters (*bottom*). The cytokine-receptor pairs that are expressed in at least 40% of CTC-neutrophil clusters are shown in *red*. **c,** The plot shows the mean 4T1-*GFP* cell number upon starvation and stimulation with IL6, IL1β, TNFα, OSM or all four cytokines together (cytokine pool). n=3; error bars represent S.E.M.; *P< 0.05 by two-sided Student's t test. **d,** Plots showing the mean percentage of Ki67-positive disseminated tumor cells (DTCs) in the bone marrow of injected mice. n=3 for all; error bars represent S.E.M.; *P<0.05 by two-sided Student's t test. **e,** The plots show normalized bioluminescence signal from the lungs of injected mice. n=4 for all; error bars represent S.E.M; *P<0.05 by two-sided Student's t test. **f,** Tumor growth curves of NSG mice injected with 4T1-Cas9-*GFP* cells expressing a control vector (Ctrl sgRNA) or sgRNAs targeting *Il1r1* or *Il6st*. n=3; Error bars represent S.E.M.; ns= not significant by two-sided Student's t test. **g,** Pie charts displaying the mean percentage of single CTCs (*grey*), CTC clusters (*green*) and CTC-neutrophil clusters (*gold*) of injected mice (*left*); n=3. The plots show the mean fold change of CTC ratios from injected mice (*right*); n=3; error bars represent S.E.M.; ns= not significant by two-sided Student's t test. **h,** The plot shows the mean percent of Ki67-positive CTCs from injected mice. n=3; error bars represent S.E.M.; *P=0.001 by two-sided Student's t test.

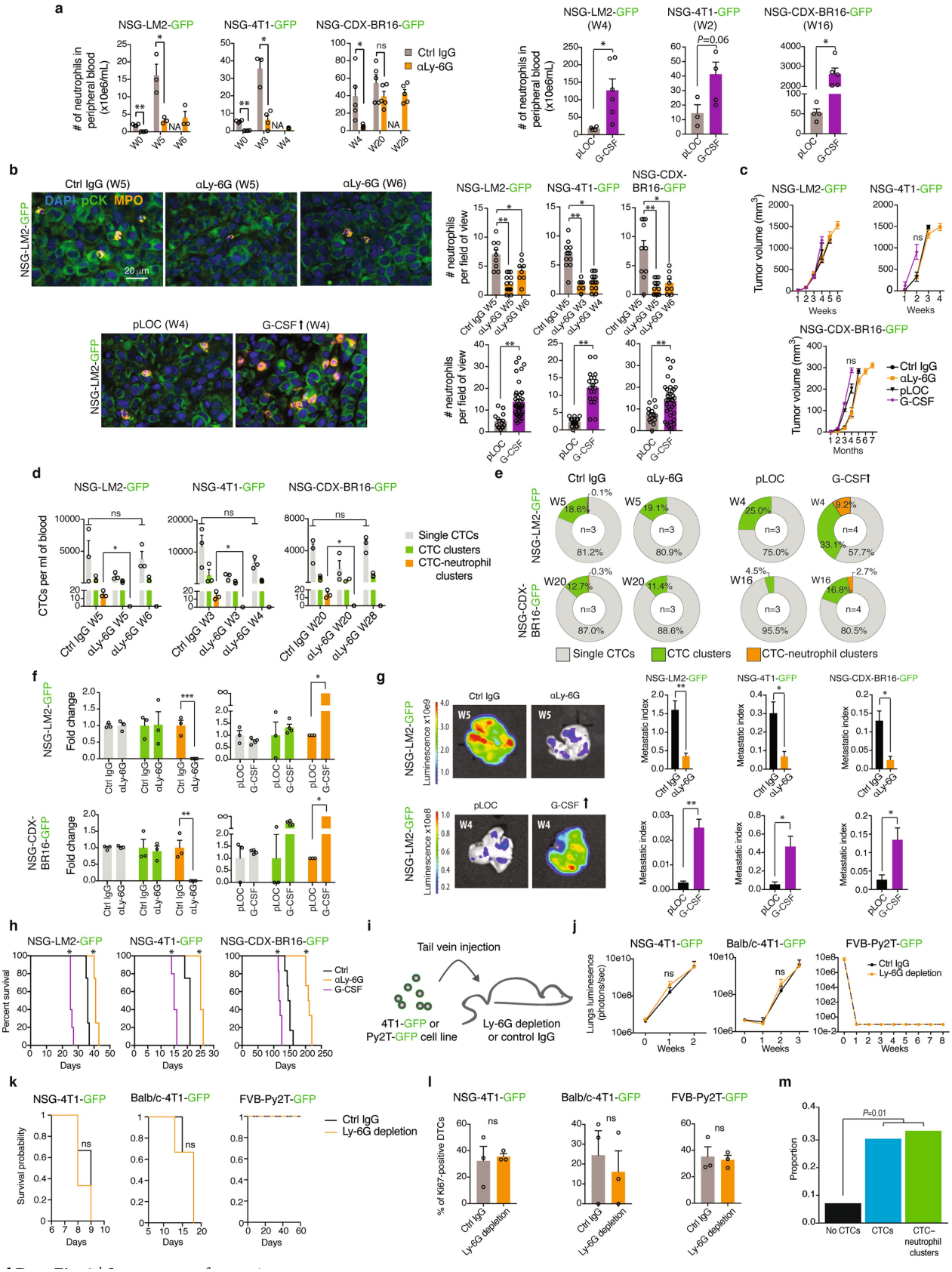


Extended Data Figure 7. Mutation analysis of single-cell whole exome sequencing data. **a,** Somatic mutation rate (mutations/Mb) of CTCs from CTC-neutrophil clusters (*n*=14) versus CTCs alone (*n*=56), normalized by donor. Lines within the violin plots show the 25th, 50th and 75th percentile, respectively, while dots represent

individual CTCs. *P* value by two-sided Wilcoxon sign-ranked test is shown. **b,** Somatic mutation rate (mutations/Mb) in all CTCs isolated from donors with CTC-neutrophil clusters (Donors (+); *n*=6) and donors without CTC-neutrophil clusters (Donors (-); *n*=5). Lines within the violin plots show the 25th, 50th and 75th percentile, respectively, while dots represent individual CTCs. ns= not significant by two-sided Wilcoxon sign-ranked test. **c,** Nucleotide substitution pattern among putative somatic mutations in CTCs isolated from donors with CTC-neutrophil clusters (Donors (+)) versus donors without CTC-neutrophil clusters (Donors (-)). *n*=6 for Donors (+) and *n*=5 for Donors (-). Lines within the violin plots show the 25th, 50th and 75th percentile, respectively, while dots represent individual CTCs. *P* value by two-sided Wilcoxon sign-ranked test is shown. **d,** The bar plots show the nucleotide context of given mutations in CTCs alone versus CTCs from CTC-neutrophil clusters. **e,** Plot showing the age distribution of Donors (+) (*n*=10) and Donors (-) (*n*=24). The *red* lines represent the mean. ns= not significant by two-sided Student's *t* test. **f,** The tile plot represent genes (*columns*) containing predicted high-impact mutations in at least two Donors (+) and in none of the Donors (-). **g,** The plots show the mean fold change for *MERTK* and *TLE1* (WT or mutated) transcripts compared to control (Ctrl) cells. *n*=3; error bars represent S.E.M.; **P*<0.004 by two-sided Student's *t* test. **h,** Tumor growth curves representing mean tumor volume measurements of NSG mice injected with 4T1 cells carrying an empty vector (pLOC), WT or mutated *MERTK* (*left*) or *TLE1* (*right*). *n*=3; error bars represent S.E.M. ns= not significant by two-sided Student's *t* test. **i,** Representative images of the primary tumor of injected mice, stained for pan cytokeratin (pCK, *green*), myeloperoxidase (MPO, *gold*) and DAPI (nuclei, *blue*) (*top*); *n*=3. The plot shows the mean number of infiltrated neutrophils per field of view within the primary tumor (*bottom*). Error bars represent S.E.M. *n*=3; ns= not significant by two-sided Student's *t* test. **j,** Representative images of the primary tumor of injected mice, stained for pan cytokeratin (pCK, *green*), myeloperoxidase (MPO, *gold*) and DAPI (nuclei, *blue*) (*top*); *n*=3. The plot shows the mean number of infiltrated neutrophils per field of view within the primary tumor (*bottom*). Error bars represent S.E.M. *n*=3; **P*=0.002 ***P*=0.0007 by two-sided Student's *t* test. **k,** Pie charts displaying the mean percentage of single CTCs (*grey*), CTC clusters (*green*) and CTC-neutrophil clusters (*gold*) in injected mice. The number of independent biological replicates (*n*) is shown for each condition.

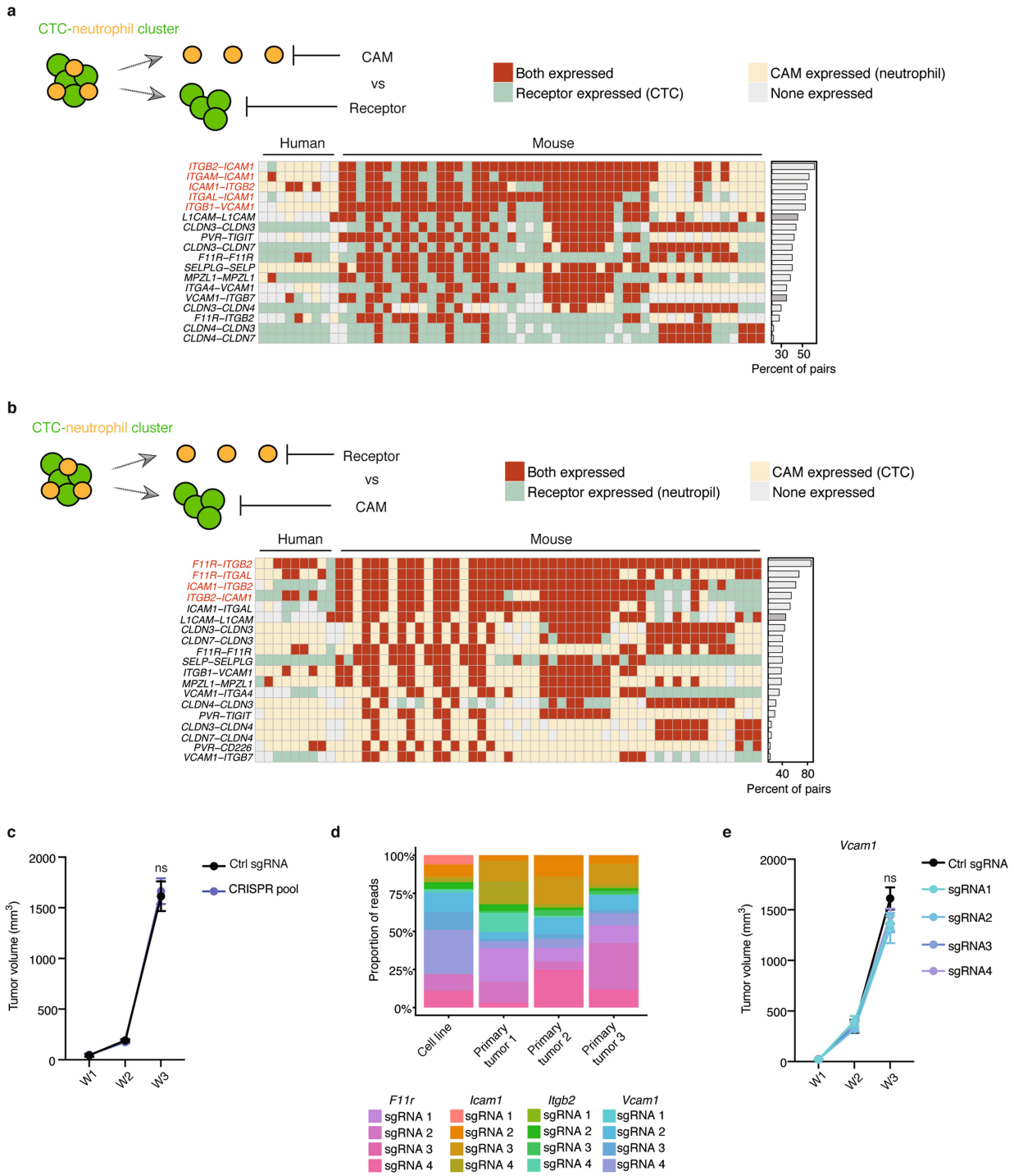


Extended Data Figure 8. Co-culture of cancer cells and neutrophils does not lead to the accumulation of key mutational events. a, Schematic of the experimental design. Neutrophils were purified from healthy donor blood and cultured with either CTC-derived cell lines (BR16, Brx50) or LM2 cells for 72h. Tumor cells were harvested, and isolated gDNA was processed for whole exome sequencing (WES). **b**, Tile plot showing the mutation status of all key loci found mutated in patients with CTC-neutrophil clusters. None of the CTC-neutrophil clusters-associated mutations were detected upon co-culture of cancer cells with neutrophils.



Extended Data Figure 9. Effects of neutrophil depletion or augmentation in mice. **a,** The plots show the mean number of neutrophils in the circulation of mice treated with Ly-6G neutralizing antibodies (αLy-6G) (*left*), or

carrying G-CSF overexpressing tumors (*right*). Error bars represent S.E.M.; the number of independent biological replicates (*n*) is provided for simplicity directly within the source data. NA=not available; **P*<0.03, ***P*<0.0001 by two-sided Student's *t* test. **b,** Representative images of the primary tumor of NSG-LM2-*GFP* mice stained for pan cytokeratin (pCK, *green*), myeloperoxidase (MPO, *gold*) and DAPI (nuclei, *blue*) (*left*); *n*=3. The plots show the mean number of infiltrated neutrophils per field of view within the tumor (*right*). W= weeks upon tumor development. Error bars represent S.E.M. *n*=3, **P*<0.03, ***P*<0.0001 by two-sided Student's *t* test. **c,** Tumor growth curves representing mean tumor volume measurements in the presence or absence of αLy-6G antibodies or G-CSF overexpression. Error bars represent S.E.M. The number of independent biological replicates (*n*) is provided for simplicity directly within the source data. ns= not significant by two-sided Student's *t* test. **d,** The plots show the mean counts of single CTCs, CTC clusters and CTC-neutrophil clusters in mice. Error bars represent S.E.M. The number of independent biological replicates (*n*) is provided for simplicity directly within the source data. ns= not significant; ND=not detected; **P*<0.05 by two-sided Student's *t* test. **e,** Pie charts displaying the mean percentage of single CTCs (*grey*), CTC clusters (*green*) and CTC-neutrophil clusters (*gold*) in NSG-LM2-*GFP* and NSG-CDX-BR16-*GFP* mice treated with αLy-6G antibodies or G-CSF overexpression. W= weeks upon tumor development; the number of independent biological replicates (*n*) is shown for each condition. **f,** The plots show the mean fold change of CTC ratios from NSG-LM2-*GFP* and NSG-CDX-BR16-*GFP* mice treated with αLy-6G antibodies or G-CSF overexpression. Error bars represent S.E.M. The number of independent biological replicates (*n*) is provided for simplicity directly within the source data. **P*=0.045, ***P*=0.01, ****P*=0.004 by two-sided Student's *t* test. **g,** Representative bioluminescence images of lungs from mice treated with αLy-6G antibodies or G-CSF overexpression (*left*); the number of independent biological replicates (*n*) is provided for simplicity directly within the source data. W = weeks upon tumor development. The plots show the mean metastatic index of mice treated with αLy-6G antibodies or G-CSF overexpression (*right*). The number of independent biological replicates (*n*) is provided for simplicity directly within the source data; error bars represent S.E.M.; *P<0.03 **P<0.01 by two-sided Student's *t* test. **h,** Kaplan-Meier survival plots showing overall survival rates of mice. The number of independent biological replicates (*n*) is provided for simplicity directly within the source data; **P*<0.02 by two-sided Log-rank test. **i,** Schematic of the experiment. NSG, FVB and Balb/c mice were pre-treated with αLy-6G antibodies or control IgG. 4T1-*GFP* cells or Py2T-*GFP* cells were then injected into the tail vein to assess metastasis development. **j,** The plots show mean normalized bioluminescence signal in the lungs of injected mice. *n*=3; error bars represent S.E.M.; ns= not significant by two-sided Student's *t* test. **k,** Kaplan-Meier survival plot of injected mice. *n*=3; ns= not significant by two-sided Student's *t* test. **l,** The plots show the mean percentage of Ki67-positive disseminated tumor cells (DTCs) collected from the bone marrow of injected mice. *n*=3; error bars represent S.E.M.; ns= not significant by two-sided Student's *t* test. **m,** The bar graph shows the proportion of breast cancer patients that were treated with G-CSF, related to their CTC status. *n*=42 for no CTCs, *n*=23 for CTCs, *n*=9 for CTC-neutrophil clusters. *P* value by two-sided Fisher's exact test is shown.



Extended Data Figure 10. Expression of cell-adhesion molecules (CAMs)-receptor pairs on CTC-neutrophil clusters. a, Schematic of the experimental design (*top*). The heatmap shows the expression landscape of cell-adhesion molecules (CAMs) and corresponding receptors that are expressed in at least 20% of CTC-neutrophil clusters (*bottom*). The CAM-receptor pairs that are expressed in at least 50% of CTC-neutrophil clusters are shown in *red*. **b,** Schematic of the experiment (*top*). The heatmap shows the expression landscape of CAM receptors and corresponding CAMs that are expressed in at least 20% of CTC-neutrophil clusters (*bottom*). The CAM-receptor

pairs that are expressed in at least 50% of CTC-neutrophil clusters are shown in *red*. **c,** Tumor growth curves representing mean tumor volume measurements of mice injected with 4T1-Cas9-*GFP* cells expressing a control vector (CTRL sgRNA) or sgRNA pools targeting *F11r*, *Icam1*, I*tgb2* and *Vcam1* (CRISPR pool). *n*=3; error bars represent S.E.M.; ns= not significant by two-sided Student's *t* test. **d,** The plot shows the proportion of reads derived from sgRNAs targeting *F11r*, *Icam1*, I*tgb2* and *Vcam1* (4 sgRNAs each) in the 4T1-Cas9-*GFP* cell line upon library transduction as well as in three primary tumors from NSG-4T1-Cas9-*GFP* mice. All sgRNAs were represented in the tumor until the end of the experiment. **e,** Tumor growth curves representing mean tumor volume measurements of mice injected with 4T1-Cas9-*GFP* cells expressing a control vector (CTRL sgRNA) or individual sgRNAs targeting *Vcam1*. *n*=3; error bars represent S.E.M.; ns= not significant by two-sided Student's *t* test.

5.1.2. "DENOSUMAB TREATMENT IS ASSOCIATED WITH THE ABSENCE OF CIRCULATING TUMOR CELLS IN PATIENTS WITH BREAST CANCER" [88]

Circulating tumor cells (CTCs) are derivatives of solid tumor lesions that detach from the tumor and enter the bloodstream [26]. In patients with breast cancer, CTCs have been shown to be predictive of a shorter disease-free survival and overall survival, with a worse prognosis in patients that present with a count of at least five CTCs per 7.5ml of blood [76, 89]. Generally, high CTC counts have been associated to a poor prognosis in multiple settings, including those patients that were newly diagnosed with metastatic breast cancer and were about to start a therapy [76, 90]. At the morphological level, breast CTCs occur in the blood of patients as single CTCs or as CTC clusters, with the latter being associated with a shorter metastasis-free survival, compared to patients in whom only single CTCs are found [21]. While the association between CTCs and bad prognosis is well established in breast cancer, CTCs are only detectable in a subset (~20-40%) of patients [21, 89]. To date, no parameters have been found that could explain CTC abundance in patients, leading to difficulties in enabling patient stratification prior to CTC-related investigations [68, 91], as well as limiting our understanding of those factors that may influence the spread of cancer. In this study, we aimed to investigate a number of clinicopathological variables, blood counts at the time of CTC isolation and detailed treatment history prior to blood sampling in a cohort of 73 consecutive patients with invasive breast cancer characterized by progressive disease, high tumor load and treatment discontinuation (or without any pre-treatment) at the time of CTC isolation, before the next line of therapy. Additionally, we did not only investigate parameters that are associated with the presence of CTCs, but we also specifically interrogated our datasets to identify features that are associated with CTC clusters. The rationale of our study was therefore to identify, in an unbiased manner (i.e. not driven by pre-existing hypotheses), clinical parameters that correlate with CTC presence in patients with progressive breast cancer.

Given previously reported correlations between number of CTCs and tumor load [92], and the findings that CTC counts predict poor prognosis in breast cancer [76, 90, 93], we focused on a group of 73 consecutive patients with invasive breast cancer that featured the following characteristics: 1) high tumor load, 2) detailed treatment history available, 3) progressive disease associated with treatment discontinuation at the time of CTC isolation (before the next line of

therapy) and 4) availability of comprehensive blood counts performed at CTC collection. Selected patients were ranging from 36 to 85 years of age and carrying either invasive ductal, lobular or inflammatory carcinoma, with a broad expression range of estrogen receptor (ER), progesterone receptor (PR), HER2 and Ki67 protein levels, as well as tumor grade varying from 1-3. Detailed characteristics for patients, therapies and statistical tests used for the analysis are listed in Extended Data Tables 1-3. Blood samples were drawn in EDTA vacutainers and processed directly with the Parsortix microfluidic device [74], with a dedicated protocol enabling the isolation of >99% of breast CTCs from unlabeled blood samples (Extended Data Figure 1A). Upon enrichment, CTCs were stained with an antibody cocktail against EpCAM, EGFR and HER2, and counterstained for the while blood cell marker CD45 (Extended Data Figure 1A and 1B). With this approach, we detected at least one CTC per 7.5ml of peripheral blood in 34 (46.6%) patients. Among these, we observed that 22 (64.7%) patients were characterized by the presence of single CTCs and that 12 (35.3%) patients had both single CTCs and CTC clusters (Extended Data Figure 1C).

We investigated a number of clinicopathological variables to identify features associated to patients in whom either single CTCs or CTC-clusters were found, compared to patients with no detectable CTCs. We first observed that previous treatments with targeted therapy (including but not limited to hormonal, anti-HER2, anti-CDK4/6 treatments), chemotherapy or radiotherapy did not correlate with the presence of CTCs (Extended Data Table 4). Yet, we found that treatment with the anti-bone resorption antibody Denosumab (received by 21/73 patients) was associated with the absence of CTCs (OR, 0.25; 95%CI, 0.06-0.86; P= 0.019). Namely, the prevalence of CTCs was 14.7% (5/21) among treated patients with Denosumab and 55.8% (29/52) among non-treated (Table 1). Further, when considering only those patients in whom CTCs were detected, the average CTC number was 9.8 for patients treated with Denosumab (n=5) versus 24.79 for nontreated patients (n=29). Despite its role as anti-bone resorption agent, the same association was not seen for bisphosphonates (P=0.784). Importantly, anti-bone resorption treatment with either Denosumab or bisphosphonates was decided based on treatment initiation date (Denosumab was approved in Switzerland in December 2011, and given as the preferred treatment option to eligible patients after that date), while patients that started receiving bisphosphonates (i.e. prior to December 2011) continued receiving bisphosphonates unless major side effects occurred (Extended Data Table 5). When restricting the analysis to the 44 patients with bone metastasis, Denosumab was administered to 20/44 of

them, and it also correlated with a reduction in CTC number compared to the remaining 24 patients with bone metastasis but no Denosumab treatment (OR, 0.22; 95%CI, 0.04-0.96; P= 0.03) (Table 2). These results were confirmed using logistic regression adjusting by age at primary diagnosis, tumor stage at diagnosis, tumor grade and histologic subtype (OR, 0.25; 95%CI, 0.06-0.82; P= 0.03). When comparing clinicopathological variables in patients that were treated or not with Denosumab, as expected, we observed a correlation with bone metastasis (OR, 22.53; 95%CI, 3.14-995.64; P= 5.6e-05; adjusted P= 0.01) (Extended Data Table 6), yet no effect in progression-free survival (Extended Data Figure 2). Together, our data show that Denosumab treatment is associated with a marked reduction of CTC counts in breast cancer patients.

	No CTC (n = 39)	CTC (n = 34)	P	Estimate (95%CI)
Age at primary diagnosis, mean (SD)	58.38 (11.85)	55.1 (11.04)	0.444	-2.76 (-8.49 - 2.84)
Age at first CTC evaluation, mean (SD)	63.53 (11.69)	59.58 (10.7)	0.163	-4.11 (-9.27 - 1.63)
Stage at Diagnosis			0.679	-
I (%)	4 (10.53%)	5 (14.71%)		
IA (%)	1 (2.63%)	0 (0%)		
II (%)	5 (13.16%)	4 (11.76%)		
IIA (%)	1 (2.63%)	4 (11.76%)		
III (%)	11 (28.95%)	7 (20.59%)		
IIIA (%)	2 (5.26%)	0 (0%)		
IIIC (%)	0 (0%)	2 (5.88%)		
IV (%)	14 (36.84%)	11 (32.35%)		
Lymphocyte node involvement			0.881	-
N0 (%)	11 (31.43%)	8 (25.81%)		
N1 (%)	11 (31.43%)	14 (45.16%)		
N2 (%)	6 (17.14%)	0 (0%)		
N3 (%)	6 (17.14%)	9 (29.03%)		
Histologic subtype			0.964	-
Invasive lobular (%)	6 (15.38%)	4 (11.76%)		
Invasive ductal (%)	31 (79.49%)	29 (85.29%)		
Inflammatory invasive lobular (%)	1 (2.56%)	1 (2.94%)		
Inflammatory (%)	2 (5.13%)	2 (5.88%)		
% of ER+ cells, mean (SD)	65.82 (42.48)	59.79 (40.93)	0.386	0 (-10 - 0)
% of PR+ cells, mean (SD)	38.97 (40.04)	26.35 (33.78)	0.171	-4 (-20 - 0)
% of KI67+ cells, mean (SD)	27.08 (17.13)	31.59 (21.55)	0.514	5 (-5 - 10)
HER2+ (%)	7 (17.95%)	7 (22.58%)	0.766	1.33 (0.35 - 5.11)
Triple - (%)	4 (10.81%)	3 (9.68%)	1.000	0.89 (0.12 - 5.73)
Tumor grade			0.985	-
1 (%)	4 (10.53%)	4 (12.12%)		
2 (%)	17 (44.74%)	14 (42.42%)		
3 (%)	17 (44.74%)	15 (45.45%)		
Bisphosphonates (%)	9 (23.68%)	7 (20.59%)	0.784	0.84 (0.23 - 2.94)
Denosumab (%)	16 (41.03%)	5 (14.71%)	0.019	0.25 (0.06 - 0.86)
Radiotherapy (%)	22 (56.41%)	13 (38.24%)	0.160	0.48 (0.17 - 1.35)
Relapse				
Any (%)	31 (79.49%)	25 (73.53%)	0.589	0.72 (0.21 - 2.45)
Local (%)	3 (7.69%)	4 (11.76%)	0.698	1.59 (0.25 - 11.72)
Metastasis (%)	26 (66.67%)	19 (55.88%)	0.470	0.64 (0.22 - 1.82)
Days between primary diagnosis and relapse, mean (SD)	1954.08 (2042.25)	1893.48 (1853.86)	0.966	-9.37 (-1000 - 756)
Established metastatic disease at CTC evaluation (%)	35 (89.74%)	30 (88.24%)	1.000	0.86 (0.15 - 5.04)
Number of metastatic sites, mean (SD)	2.09 (1.01)	1.9 (0.94)	0.473	0 (-1 - 0)
Metastasis site				
Bone (%)	27 (69.23%)	17 (51.52%)	0.150	0.45 (0.15 - 1.28)
Liver (%)	10 (25.64%)	12 (36.36%)	0.447	1.57 (0.51 - 4.9)
Lymphnode (%)	9 (23.08%)	10 (30.3%)	0.599	1.38 (0.43 - 4.54)
Pleural (%)	7 (17.95%)	2 (6.06%)	0.162	0.29 (0.03 - 1.68)
Peritoneal (%)	3 (7.69%)	4 (12.12%)	0.698	1.59 (0.25 - 11.72)
Lung (%)	4 (10.26%)	4 (12.12%)	1.000	1.16 (0.2 - 6.83)
Skin (%)	3 (7.69%)	0 (0%)	0.243	0 (0 - 2.74)
Brain (%)	2 (5.13%)	2 (6.06%)	1.000	1.15 (0.08 - 16.76)
Uterus (%)	1 (2.56%)	1 (3.03%)	1.000	1.15 (0.01 - 92.67)
Muscular (%)	1 (2.56%)	2 (6.06%)	0.595	2.35 (0.12 - 143.61)

Table 1. Clinical Features of Patients with Circulating Tumor Cells. The table shows clinical features of patients with and without circulating tumor cells (CTCs). Abbreviations: ER, estrogen receptor; HER2, human epidermal growth factor receptor 2; PR, progesterone receptor; SD, standard deviation.

	Number of samples	No CTCs	CTCs	P	Estimate (95%CI)	P	Estimate (95%CI)
Reference	28	12 (43%)	16 (57%)	ref		-	-
Bone metastasis	24	11 (46%)	13 (54%)	1.000	0.89 (0.26-3.05)	ref	
Bone metastasis and Denosumab	20	16 (80%)	4 (20%)	0.017	0.19 (0.04-0.81)	0.030	0.22 (0.04-0.96)

Table 2. Circulating Tumor Cells Detection According to Denosumab Treatment and Bone Metastasis. The table shows the number of patients with and without circulating tumor cells (CTCs) among individuals with bone metastasis that were treated or not with Denosumab.

We further asked whether any clinicopathological variables might be associated specifically to the presence of CTC clusters, compared to patients in whom CTC clusters were not found (i.e. having either single CTCs or no CTCs). We found that both younger age at primary diagnosis as well as younger age at first CTC evaluation were associated to the presence of CTC clusters (Table 3). Particularly, we observed an average age at primary diagnosis of 50.63 years (SD, 12.60) for patients with CTC clusters and 58.08 years (SD, 10.99) for patients with no CTC clusters (P= 0.033), as well as a average age at first CTC evaluation of 54.87 years (SD, 12.14) for patients with CTC clusters and 63.03 years (SD, 10.77) for patients with no CTC clusters (P=0.025). We also observed that while HER2 is expressed in 22% (13/61) of patients with no CTC clusters (7/39, i.e. 17.95% of patients with no CTCs and 6/22, i.e. 30% of patients with single CTCs only), it is only expressed in one patient (9.09%) with CTC clusters (OR, 0.36; 95% CI, 0.01-2.97; P = 0.44); even though the relationship between HER2 negativity and CTC clusters did not reach statistical significance (Table 3). When considering only patients with CTCs, and comparing those with CTC clusters versus those with single CTCs, we found that also in this context younger age at primary diagnosis (P= 0.044) and younger age at first CTC evaluation (P= 0.058) are associated with the presence of CTC-clusters (Extended Data Table 7).

Additionally, to investigating the clinical parameters summarized above, for each patient we also evaluated comprehensive blood counts performed at CTC collection. We first asked whether blood-related parameters were associated with the presence of CTCs (either single or clustered), compared to patients in whom CTCs were not detected. We observed that patients with detectable CTCs had a lower red blood cell (RBC) count (OR, -0.42; 95% CI, -0.8 - -0.08; P= 0.019) compared to patients with no CTCs (Table 4).

	No CTC clusters (n = 61)	CTC clusters (n = 12)	P	Estimate (95%CI)
Age at primary diagnosis, mean (SD)	58.08 (10.99)	50.63 (12.6)	0.033	-8.26 (-15.3 - -0.44)
Age at first CTC evaluation, mean (SD)	63.03 (10.77)	54.87 (12.14)	0.025	-8.3 (-16.06 - -1.04)
Stage at Diagnosis			0.726	-
I (%)	7 (11.67%)	2 (16.67%)		
IA (%)	1 (1.67%)	0 (0%)		
II (%)	7 (11.67%)	2 (16.67%)		
IIA (%)	4 (6.67%)	1 (8.33%)		
III (%)	16 (26.67%)	2 (16.67%)		
IIIA (%)	2 (3.33%)	0 (0%)		
IIIC (%)	1 (1.67%)	1 (8.33%)		
IV (%)	21 (35%)	4 (33.33%)		
Lymphocyte node involvement			0.855	-
N0 (%)	15 (27.27%)	4 (36.36%)		
N1 (%)	22 (40%)	3 (27.27%)		
N2 (%)	6 (10.91%)	0 (0%)		
N3 (%)	11 (20%)	4 (36.36%)		
Histologic subtype			0.679	-
Invasive lobular (%)	9 (14.75%)	1 (8.33%)		
Invasive ductal (%)	49 (80.33%)	11 (91.67%)		
Inflammatory invasive lobular (%)	1 (1.64%)	1 (8.33%)		
Inflammatory (%)	3 (4.92%)	1 (8.33%)		
% of ER+ cells, mean (SD)	62.34 (41.8)	66.42 (42.1)	0.675	0 (-10 - 20)
% of PR+ cells, mean (SD)	32.71 (37.79)	34 (37.3)	0.888	0 (-10 - 20)
% of KI67+ cells, mean (SD)	30 (19.65)	23 (16.43)	0.384	-5 (-20 - 10)
HER2+ (%)	13 (22.03%)	1 (9.09%)	0.442	0.36 (0.01 - 2.97)
Triple - (%)	7 (12.28%)	0 (0%)	0.588	0 (0 - 3.7)
Tumor grade			0.093	-
1 (%)	5 (8.33%)	3 (27.27%)		
2 (%)	26 (43.33%)	5 (45.45%)		
3 (%)	29 (48.33%)	3 (27.27%)		
Bisphosphonates (%)	14 (23.33%)	2 (16.67%)	1.000	0.66 (0.06 - 3.68)
Denosumab (%)	19 (31.15%)	2 (16.67%)	0.489	0.45 (0.04 - 2.41)
Radiotherapy (%)	30 (49.18%)	5 (41.67%)	0.756	0.74 (0.17 - 3.06)
Relapse				
Any (%)	47 (77.05%)	9 (75%)	1.000	0.9 (0.19 - 5.83)
Local (%)	4 (6.56%)	3 (25%)	0.082	4.61 (0.58 - 32.62)
Metastasis (%)	40 (65.57%)	5 (41.67%)	0.193	0.38 (0.08 - 1.59)
Days between primary diagnosis and relapse, mean (SD)	1969.49 (2003.96)	1636.67 (1538.48)	0.633	-236.86 (-1643 - 1203)
Established metastatic disease at CTC evaluation (%)	54 (88.52%)	11 (91.67%)	1.000	1.42 (0.15 - 70.01)
Number of metastatic sites, mean (SD)	1.96 (0.98)	2.18 (0.98)	0.452	0 (0 - 1)
Metastasis site				
Bone (%)	37 (61.67%)	7 (58.33%)	1.000	0.91 (0.22 - 4.08)
Liver (%)	19 (31.67%)	3 (25%)	1.000	0.74 (0.12 - 3.42)
Lymphnode (%)	15 (25%)	4 (33.33%)	0.497	1.52 (0.29 - 6.74)
Pleural (%)	9 (15%)	0 (0%)	0.339	0 (0 - 2.57)
Peritoneal (%)	5 (8.33%)	2 (16.67%)	0.323	2.21 (0.19 - 16.05)
Lung (%)	7 (11.67%)	1 (8.33%)	1.000	0.7 (0.01 - 6.47)
Skin (%)	3 (5%)	0 (0%)	1.000	0 (0 - 12.81)
Brain (%)	3 (5%)	1 (8.33%)	0.521	1.74 (0.03 - 24.14)
Uterus (%)	1 (1.67%)	1 (8.33%)	0.304	5.27 (0.06 - 433.34)
Muscular (%)	2 (3.33%)	1 (8.33%)	0.421	2.63 (0.04 - 54.78)

Table 3. Clinical Features of Patients with Circulating Tumor Cell Clusters. The table shows clinical features of patients with and without circulating tumor cell clusters (CTC clusters). Abbreviations: ER, estrogen receptor; HER2, human epidermal growth factor receptor 2; PR, progesterone receptor; SD, standard deviation.

	No CTC (n = 39)	CTC (n = 34)	P	Estimate (95%CI)
CA 15-3, mean (SD)	223.71 (384.68)	1084.15 (4136.87)	0.658	6.7 (-19.2 - 87.6)
Alkaline phosphatase, mean (SD)	105.47 (103.98)	198.15 (365.58)	0.401	6 (-12 - 27)
Calcium (korr), mean (SD)	2.34 (0.15)	2.32 (0.25)	0.145	-0.06 (-0.13 - 0.02)
CRP, mean (SD)	31.92 (47.56)	26.87 (47.69)	0.982	0 (-8.8 - 3.8)
LDH, mean (SD)	281.61 (118.18)	300.15 (228.57)	0.772	-5 (-36 - 23)
RBC 10^12/L, mean (SD)	4.37 (0.56)	3.85 (0.77)	0.019	-0.42 (-0.8 - -0.08)
HGB g/L, mean (SD)	130.14 (19.85)	118.15 (24.11)	0.051	-11 (-21 - 0)
HCT L/L, mean (SD)	0.38 (0.06)	0.35 (0.06)	0.053	-0.03 (-0.06 - 0)
MCV fL, mean (SD)	87.46 (5.68)	89.73 (5.26)	0.227	2 (-1 - 4)
MCH pg, mean (SD)	29.62 (2.52)	30.66 (1.96)	0.157	0.8 (-0.3 - 1.8)
MCHC g/L, mean (SD)	339.03 (13.98)	341.19 (12.79)	0.667	2 (-6 - 8)
WBC 10^9/L, mean (SD)	7.35 (2.13)	7.24 (3.53)	0.334	-0.6 (-1.88 - 0.81)
Neutrophile 10^9/L, mean (SD)	5.33 (1.87)	5.12 (2.87)	0.239	-0.63 (-1.65 - 0.56)
Lymphocyte10^9/L, mean (SD)	1.37 (0.67)	1.41 (0.85)	0.941	-0.02 (-0.38 - 0.38)
Monocyte 10^9/L, mean (SD)	0.42 (0.12)	0.44 (0.2)	0.843	-0.01 (-0.08 - 0.07)
Eosinophil 10^9/L, mean (SD)	0.16 (0.18)	0.17 (0.12)	0.423	0.02 (-0.03 - 0.09)
Basophil 10^9/L, mean (SD)	0.04 (0.07)	0.04 (0.03)	0.256	0.01 (-0.01 - 0.02)
LUC 10^9/L, mean (SD)	0.15 (0.22)	0.12 (0.08)	0.912	0 (-0.03 - 0.02)
PLT 10^9/L, mean (SD)	289.92 (139.86)	249.93 (91.52)	0.277	-30 (-84 - 24)
MPV fL, mean (SD)	8.2 (1.51)	8.68 (1.52)	0.109	0 (0 - 1)

Table 4. Complete Blood Counts in Patients with Circulating Tumor Cells. The table shows complete blood counts in patients with and without circulating tumor cells (CTCs). Abbreviations: CA 15-3, cancer antigen 15-3; CRP, C-reactive protein; HCT, hematocrit; HGB, hemoglobin; LDH, lactate dehydrogenase; LUC, large unstained cells; MCH, mean corpuscular hemoglobin; MCHC, mean corpuscular hemoglobin concentration; MCV, mean corpuscular volume; MPV, Mean platelet volume; PLT, platelets; RBC, red blood cells; WBC, white blood cells.

We then asked whether specific blood-related parameters could be associated to the presence of CTC clusters, compared to patients with no CTC clusters (i.e. having either no CTCs or single CTCs only). In this case, we found that patients with CTC 13/23 clusters have 14-fold higher levels of CA 15-3 tumor marker (P=0.021), higher mean corpuscular volume (MCV) (P=0.033), higher white blood cells counts (WBC) (P=0.03) and higher mean platelet volume (MPV) (P=0.032) compared to patients in whom CTC clusters are not found (Table 5). We also restricted this analysis to patients with CTCs and compared patients with CTC clusters

to patients with only single CTCs. In this setting, we further confirmed that patients with CTC clusters have 38-fold higher CA 15-3 tumor antigen (P=0.0089), as well as nearly two-fold higher total white blood cell (WBC) counts (P=0.0045) and higher neutrophil counts (P=0.03) (Extended Data Table 8).

	No CTC clusters (n = 61)	CTC clusters (n = 12)	P	Estimate (95%CI)
CA 15-3, mean (SD)	172.5 (324.45)	2554.6 (6387.64)	0.021	204.16 (9.7 - 515)
Alkaline phosphatase, mean (SD)	106.4 (104.27)	310.25 (525.57)	0.301	10 (-12 - 74.58)
Calcium (korr), mean (SD)	2.33 (0.14)	2.36 (0.35)	0.698	-0.02 (-0.16 - 0.14)
CRP, mean (SD)	27.76 (44.47)	37.25 (58.6)	0.279	2.9 (-4 - 21.4)
LDH, mean (SD)	271 (102.59)	363.33 (329.24)	0.463	17 (-26 - 76)
RBC 10^12/L, mean (SD)	4.26 (0.6)	3.66 (0.93)	0.078	-0.51 (-1.16 - 0.06)
HGB g/L, mean (SD)	127.52 (20.18)	114.18 (29.19)	0.183	-12 (-30 - 6)
HCT L/L, mean (SD)	0.38 (0.05)	0.34 (0.08)	0.159	-0.03 (-0.09 - 0.01)
MCV fL, mean (SD)	87.7 (5.58)	91.73 (4.41)	0.033	4 (0 - 7)
MCH pg, mean (SD)	29.8 (2.42)	31.24 (1.51)	0.064	1.1 (0 - 2.3)
MCHC g/L, mean (SD)	339.8 (13.38)	340.64 (14.25)	0.646	2 (-10 - 10)
WBC 10^9/L, mean (SD)	6.87 (2.25)	9.38 (3.99)	0.030	2.54 (0.26 - 4.68)
Neutrophile 10^9/L, mean (SD)	4.94 (1.87)	6.65 (3.6)	0.177	1.22 (-0.52 - 3.68)
Lymphocyte10^9/L, mean (SD)	1.38 (0.69)	1.42 (1)	0.756	-0.11 (-0.62 - 0.6)
Monocyte 10^9/L, mean (SD)	0.41 (0.13)	0.5 (0.24)	0.336	0.06 (-0.08 - 0.26)
Eosinophil 10^9/L, mean (SD)	0.16 (0.16)	0.17 (0.14)	0.974	0 (-0.07 - 0.1)
Basophil 10^9/L, mean (SD)	0.04 (0.06)	0.04 (0.04)	0.983	0 (-0.02 - 0.02)
LUC 10^9/L, mean (SD)	0.14 (0.19)	0.12 (0.11)	0.471	-0.01 (-0.05 - 0.02)
PLT 10^9/L, mean (SD)	278.02 (125.8)	251.5 (109.99)	0.667	-15.84 (-94 - 47)
MPV fL, mean (SD)	8.32 (1.62)	8.82 (0.87)	0.032	1 (0 - 1)

Table 5. Complete Blood Counts in Patients with Circulating Tumor Cell Clusters. The table shows complete blood counts in patients with and without circulating tumor cell clusters (CTC clusters). Abbreviations: CA 15-3, cancer antigen 15-3; CRP, C-reactive protein; HCT, hematocrit; HGB, hemoglobin; LDH, lactate dehydrogenase; LUC, large unstained cells; MCH, mean corpuscular hemoglobin; MCHC, mean corpuscular hemoglobin concentration; MCV, mean corpuscular volume; MPV, Mean platelet volume; PLT, platelets; RBC, red blood cells; WBC, white blood cells.

In a selected cohort of 73 patients with progressive invasive breast cancer, we provide a detailed description of a number of clinicopathological parameters and blood counts at the time of CTC isolation that correlate with the presence of single CTCs and CTC clusters. Interestingly, we observe that treatment with the monoclonal antibody Denosumab in patients with bone metastasis strongly correlates with the absence of CTCs from their peripheral circulation, suggesting a scenario whereby the treatment itself might influence CTC spread from the bone

tissue. Importantly, this correlation is not seen for treatment with the anti-bone resorption drug bisphosphonate, possibly because of different administration routes or dosing schedules [94] or alternatively, potential off-target binding of Denosumab to proteins other than RANKL.

While focusing on clinical parameters, our study does not provide molecular insights into the mechanism of action of Denosumab in the context of its role in inhibiting CTC generation. Yet, considering that most Denosumab-treated patients are characterized by bone metastatic disease but no primary breast tumor (which has been surgically removed prior to Denosumab treatment), CTCs represent derivatives of their bone metastatic lesions. In this setting, we speculate that the effect of Denosumab in suppressing CTC generation could be the result of RANKL inhibition within the bone, preventing the maturation of pre-osteoclasts into osteoclasts [95] and protecting the bone from degradation, leading to a lower likelihood of a bone metastatic lesion to shed CTCs. However, we cannot exclude an action of Denosumab on breast cancer cells themselves, which have been previously shown to express high RANK levels [96, 97] and may be susceptible to its inhibition. Prospective studies and molecular assays will be needed to specifically dissect the role and mechanism of action of Denosumab in CTC generation.

Recently, a phase 3 clinical trial designed to determine the long-term effects of Denosumab treatment (D-CARE; NCT01077154) showed no benefits in metastasis-free survival and overall survival of breast cancer patients. While individuals within this study mainly comprised patients with early breast cancer, i.e. stage IIB to IIIC, our patient cohort was largely dominated by patients with stage IV disease. While we are not aware of CTC enumeration data being evaluated within the D-CARE study, it is possible that Denosumab might play a different role in the intravasation of bone metastasis-derived CTCs (as seen in our study) as opposed to primary tumor-derived CTCs (D-CARE).

Among other correlations, we observe an intriguing association between the absence of HER2 expression in the primary tumor and the presence of CTC clusters. While this result did not reach statistical significance, our observation regarding HER2 does not seem to be influenced by the metastatic tropism of HER2-positive breast cancers, and it might reveal important insights into the signaling networks involving CTC clusters formation, also considering HER2 expression fluctuations in CTCs and breast cancer metastasis [98, 99]. In other words, we

speculate that HER2 signaling might influence cancer cells to intravasate as single CTCs, while its absence might poise them towards collective invasion into the bloodstream. This hypothesis will require experimental testing to be addressed.

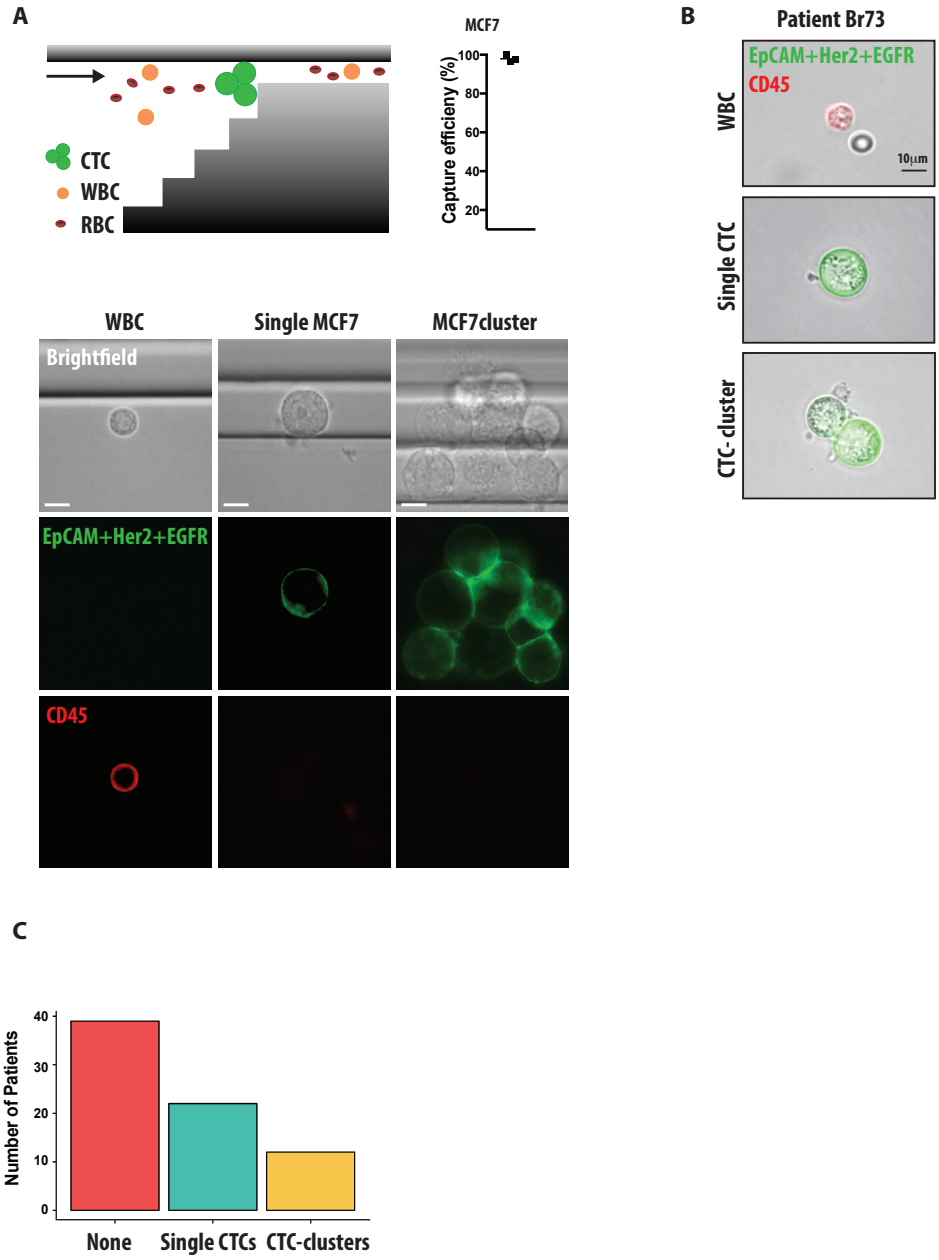
We also found that CTC clusters, but not CTCs in general, are more prevalent in younger patients. Both CTC clusters and younger age have been associated with worse prognosis and reduced survival rates [21, 89, 100-103]. In this case, it is unlikely that younger age represents an independent risk factor for CTC clusters formation, but rather it may reflect an association to tumor aggressiveness [104].

Lastly, blood counts at the time of CTC collection provide evidence for applying well-established, cost-effective and widespread blood testing strategies to stratify patients with higher likelihood to present with detectable CTCs. For instance, we find that lower RBC count is a good correlate with the presence of CTCs. Additionally, CA 15-3 tumor antigen is highly increased in patients with CTC clusters, possibly reflecting a higher tumor load but also tumors that are characterized by an elevated shedding of MUC-1-containing cells into the bloodstream [105]. A functional relationship between MUC-1 and CTC clusters remains to be investigated. We also observe that higher MCV, higher MPV and higher WBC counts correlate with the presence of CTC clusters. We envision these parameters to be used to stratify patient populations to conduct CTC-related studies in the setting of advanced breast cancer.

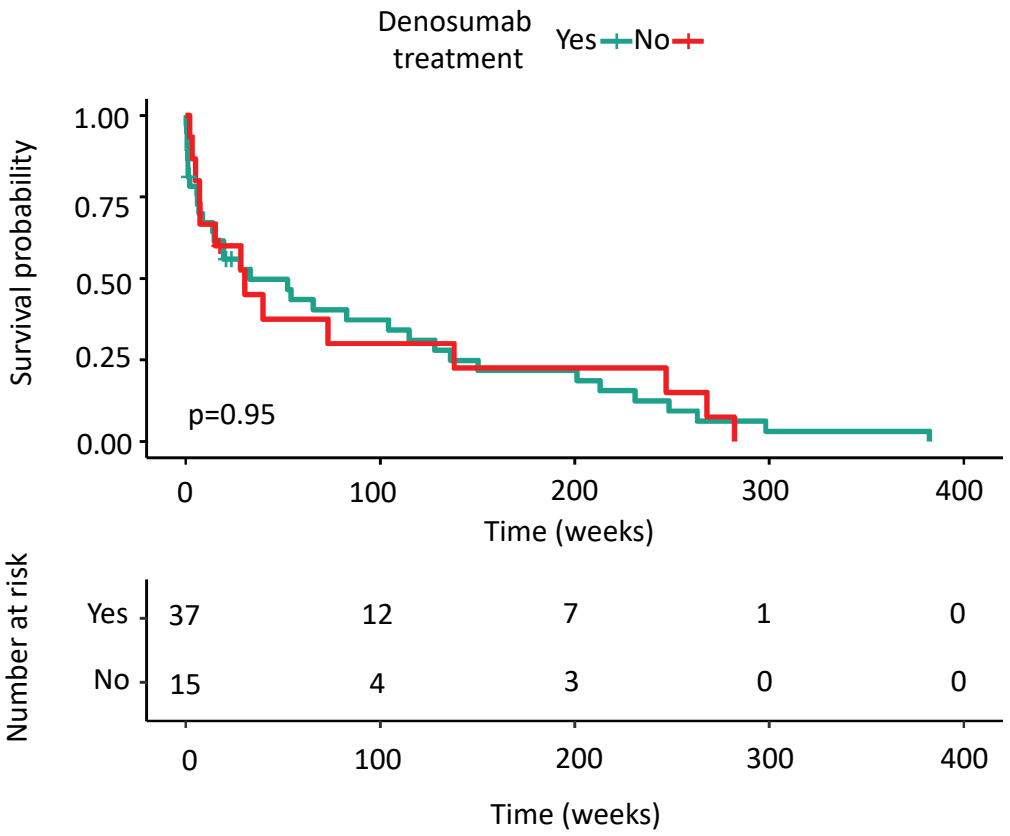
Altogether, our study is meant as an exploratory analysis to evaluate the association of multiple clinical predictors with the presence of CTCs. Given the high number of hypotheses tested and the relatively low number of patient samples enrolled in the study (n=73), none of the associations reported show a P value lower than 0.05 after adjustment for multiple comparisons, with the exception of the correlation between Denosumab treatment and the presence of bone metastasis (adjusted P = 0.01). For this reason, subsequent prospective and experimental studies should be conducted to validate the associations that are presented in this work, including the role of Denosumab in CTC shedding.

Our data provide evidence of the association between treatment with the monoclonal antibody Denosumab and the absence of CTCs from the peripheral circulation of breast cancer patients.

This finding suggests that Denosumab treatment may be beneficial to reduce cancer spread in patients that are diagnosed with bone metastasis.



Extended Data Figure 1. Circulating tumor cell (CTC) capture strategy. A. Schematic drawing showing the size-based capturing principle of the Parsortix microfluidic device (*left*). Plot showing the capture efficiency of the Parsortix microfluidic device for MCF7 cells spiked in healthy blood samples (*right*). Representative images of a captured MCF7 single cell, a cell cluster (*green*) and a contaminant white blood cell (WBC; red) in the Parsortix microfluidic cassette (*bottom*). **B.** Representative images of a captured single CTC, a CTC cluster (*green*) and a contaminant WBC (*red*) from a breast cancer patient sample. **C.** Bar graph showing the number of patients in whom no CTCs, single CTCs or CTC clusters were found.



Extended Data Figure 2. Progression-free survival of patients who were treated or not with denosumab. Kaplan-Meier curve showing the progression-free survival probability of patients who were treated (*red*) or not (*green*) with denosumab (*top*). $P = 0.95$ by pairwise log-rank test. The table shows the number of patients at each time point (*bottom*).

Drug	Action	Group	Other names
afinitor	mTOR inhibitor	Targeted therapy	afinitor
abemaciclib	CDK4 and CDK6 inhibitor	Targeted therapy	abemaciclib
adjuvant tcbx6 *taxol+carbo)	chemotherapy	Chemotherapy	adjuvant tcbx6 *taxol+carbo)
afinitor	mTOR inhibitor	Targeted therapy	everolimus;afinitor
anastrozol	aromatase inhibitor	Hormone therapy	anastrozol;arimidex
arimidex	aromatase inhibitor	Hormone therapy	anastrozol;arimidex
aromasin	aromatase inhibitor	Hormone therapy	aromasin;exemestan
atezolizumab	PD-L1 inhibitor	Immunotherapy	atezolizumab
avastin	VEGF inhibitor	Targeted therapy	avastin
caelyx	chemotherapy	Chemotherapy	caelyx;pld
carboplatin	chemotherapy	Chemotherapy	carboplatin
decapeptyl	anti-estrogen	Hormone therapy	decapeptyl
depocyte intrathekal	chemotherapy	Chemotherapy	depocyte intrathekal
ec	chemotherapy	Chemotherapy	ec
epirubicin	chemotherapy	Chemotherapy	epirubicin
eribulin	chemotherapy	Chemotherapy	halaven;eribulin
everolimus	mTOR inhibitor	Targeted therapy	everolimus;afinitor
exemestan	aromatase inhibitor	Hormone therapy	aromasin;exemestan
farlutal	Synthetic progesteron	Hormone therapy	farlutal
faslodex	anti-estrogen	Hormone therapy	fulvestrant;faslodex
femara	aromatase inhibitor	Hormone therapy	letrozol;femara
fulvestrant	anti-estrogen	Hormone therapy	faslodex;fulvestrant
gemzar	chemotherapy	Chemotherapy	gemzar
gnrh	anti-estrogen	Hormone therapy	gnrh;zoladex
halaven	chemotherapy	Chemotherapy	eribulin;halaven
herceptin	Her2 inhibitor	Targeted therapy	herceptin
il-dox	chemotherapy	Chemotherapy	il-dox
kadcyla	Her2+tubulin blocker	Targeted therapy	kadcyla;tdm1
l-pld	chemotherapy	Chemotherapy	l-pld
lapatinib	Her2+EGFR blocker	Targeted therapy	lapatinib
letrozol	aromatase inhibitor	Hormone therapy	femara;letrozol
lucrin	anti-estrogen	Hormone therapy	lucrin
nab-paclitaxel	chemotherapy	Chemotherapy	nab-paclitaxel
navelbine	chemotherapy	Chemotherapy	navelbine;vinorelbin
ofs	estrogen blocker	Hormone therapy	ofs
paclitaxel	chemotherapy	Chemotherapy	paclitaxel;taxol
palbociclib	CDK4 and CDK6 inhibitor	Targeted therapy	palbociclib
perjeta	Her2 inhibitor	Targeted therapy	perjeta
pld	chemotherapy	Chemotherapy	caelyx;pld
tam	anti-estrogen	Hormone therapy	tam;tamoxifen
tamoxifen	anti-estrogen	Hormone therapy	tam;tamoxifen
taxol	chemotherapy	Chemotherapy	paclitaxel;taxol
taxotere	chemotherapy	Chemotherapy	taxotere
tdm1	Her2+tubulin blocker	Targeted therapy	kadcyla;tdm1
vinorelbin	chemotherapy	Chemotherapy	navelbine;vinorelbin
xeloda	chemotherapy	Chemotherapy	xeloda
zoladex	anti-estrogen	Hormone therapy	gnrh;zoladex
adriamycin-cyclophosphamide	chemotherapy	Chemotherapy	adriamycin-cyclophosphamide;ac
carboplatin-docetaxel	chemotherapy	Chemotherapy	carboplatin-docetaxel;tcb
tcb	chemotherapy	Chemotherapy	carboplatin-docetaxel;tcb
cisplatin	chemotherapy	Chemotherapy	cisplatin
adriblastin	chemotherapy	Chemotherapy	adriblastin
endoxan	chemotherapy	Chemotherapy	endoxan

Extended Data Table 1. Drug classification. The table shows the drug classification used for the analysis, grouping drugs into targeted therapy, chemotherapy, hormone therapy and immunotherapy.

	Type of Variable	Statistical Test
Age at primary diagnosis, mean (SD)	Numerical	two-sided Wilcoxon rank-sum test
Age at first CTC evaluation, mean (SD)	Numerical	two-sided Wilcoxon rank-sum test
Stage at Diagnosis	Ordinal	Kruscal Wallis
I (%)		
IA (%)		
II (%)		
IIA (%)		
III (%)		
IIIA (%)		
IIIC (%)		
IV (%)		
Lymphocyte node involvement	Ordinal	Kruscal Wallis
N0 (%)		
N1 (%)		
N2 (%)		
N3 (%)		
Histologic subtype	Categorical >2-level	Fisher test
Invasive lobular (%)		
Invasive ductal (%)		
Inflammatory invasive lobular (%)		
Inflammatory (%)		
% of ER+ cells, mean (SD)	Numerical	two-sided Wilcoxon rank-sum test
% of PR+ cells, mean (SD)	Numerical	two-sided Wilcoxon rank-sum test
% of KI67+ cells, mean (SD)	Numerical	two-sided Wilcoxon rank-sum test
HER2+ (%)	Categorical 2-level	Fisher test
Triple - (%)	Categorical 2-level	Fisher test
Tumor grade	Ordinal	Kruscal Wallis
1 (%)		
2 (%)		
3 (%)		
Bisphosphonates (%)	Categorical 2-level	Fisher test
Denosumab (%)	Categorical 2-level	Fisher test
Radiotherapy (%)	Categorical 2-level	Fisher test
Relapse		
Any (%)	Categorical 2-level	Fisher test
Local (%)	Categorical 2-level	Fisher test
Metastasis (%)	Categorical 2-level	Fisher test
Days between primary diagnosis and relapse, mean (SD)	Numerical	two-sided Wilcoxon rank-sum test
Established metastatic disease at CTC evaluation (%)	Categorical 2-level	Fisher test
Number of metastatic sites, mean (SD)	Numerical	two-sided Wilcoxon rank-sum test
Metastasis site		
Bone (%)	Categorical 2-level	Fisher test
Liver (%)	Categorical 2-level	Fisher test
Lymphnode (%)	Categorical 2-level	Fisher test
Pleural (%)	Categorical 2-level	Fisher test
Peritoneal (%)	Categorical 2-level	Fisher test
Lung (%)	Categorical 2-level	Fisher test
Skin (%)	Categorical 2-level	Fisher test
Brain (%)	Categorical 2-level	Fisher test
Uterus (%)	Categorical 2-level	Fisher test
Muscular (%)	Categorical 2-level	Fisher test

Extended Data Table 2. Variable classification and statistical test applied. The table shows the type of variable and statistical test used for the analysis of individual clinicopathological parameters.

Patient	Age at primary diagnosis	Age at first CTC evaluation	% of ER+ cells	% of PR+ cells	% of KI67+ cells	HER2+	Histologic subtype	Tumor grade	Established metastatic disease at CTC evaluation	ECOG at CTC evaluation	Therapy at CTC evaluation	Accumulated Therapy at CTC evaluation
Br1	47.69	54.18	90	90	NA	0	ID	3	yes	1	none	none
Br2	57.6	72.39	1	1	NA	NA	IL	3	yes	1	Hormone	Hormone; Chemotherapy
Br3	56.77	57.54	35	0	40	1	ID	3	yes	1	Targeted; Chemotherapy	Targeted; Chemotherapy
Br4	53.48	54.96	80	5	30	0	IL	3	yes	1	Chemotherapy	Chemotherapy
Br6	74.74	81.68	70	80	NA	0	IL	2	yes	3	none	none
Br7	41.88	41.92	90	90	25	0	ID inflammatory	2	yes	1	none	none
Br8	64.51	69.5	70	30	10	0	ID	2	yes	1	Hormone	Hormone
Br9	60.16	62.67	70	30	60	0	ID	3	yes	1	Hormone	Chemotherapy ; Hormone
Br10	48.45	64.72	100	0	NA	0	ID	1	yes	0-1	Hormone	Hormone; Chemotherapy
Br11	53.81	56.31	0	0	50	0	ID	3	yes	1	Chemotherapy	Chemotherapy
Br12	60.1	67.11	90	70	NA	0	IL	2	yes	NA	Chemotherapy	Hormone; Chemotherapy ; Targeted
Br13	56.02	61.47	100	5	5	0	IL	2	yes	0-1	Targeted; Hormone	Hormone; Targeted
Br14	72.68	74.76	100	10	NA	0	ID	2	yes	1	none	Hormone; none
Br15	38.05	42.9	100	75	NA	NA	ID	1	yes	1	Hormone	Hormone; Targeted
Br16	38.71	49.04	100	75	NA	0	ID	2	yes	2	Chemotherapy	Hormone; Chemotherapy
Br17	54.82	66.94	1	1	NA	0	ID	1	yes	2	Hormone	Chemotherapy ; Hormone
Br18	59.89	63.61	80	0	NA	0	ID	3	yes	0-1	Chemotherapy	Hormone; Chemotherapy
Br19	40.05	42.48	0	0	NA	1	ID	2	yes	1	Targeted	Targeted; Chemotherapy
Br20	40.86	45.26	80	80	35	0	ID	3	yes	0-1	none	none
Br21	55.47	69.23	100	70	NA	0	IL	2	yes	1	none	none
Br22	53.35	66.15	100	90	NA	0	ID	2	yes	01.Feb	Hormone	Hormone
Br23	63.02	63.05	100	0	NA	0	IL	2	yes	1	Hormone	Hormone
Br25	39.02	45.84	70	15	NA	0	ID	2	yes	0-1	Hormone	Hormone
Br27	62.22	64.26	100	5	20	1	ID	2	yes	1	Targeted; Chemotherapy	Targeted; Chemotherapy
Br29	53.22	54.32	0	0	70	0	ID	3	yes	02.Mar	none	none
Br30	48.99	50.64	80	70	25	0	ID	3	yes	1	none	Chemotherapy ; none
Br31	81.68	81.92	100	5	25	0	ID	3	yes	1	Hormone	Hormone
Br32	68.32	68.43	90	100	20	1	ID	3	yes	NA	Hormone; Targeted	Hormone; Targeted
Br33	53.18	68.71	0	0	NA	1	ID	3	yes	1	Targeted	Targeted; Chemotherapy
Br34	49.93	65.51	1	1	NA	0	IL	2	yes	0	Hormone	Hormone
Br35	71.97	78.42	70	50	15	0	IL	2	yes	1	Chemotherapy	Chemotherapy
Br36	59.84	59.89	50	0	50	1	ID	3	no	0-1	Hormone	Hormone
Br37	67.19	67.25	70	0	27.5	1	ID	3	yes	0	Targeted; Chemotherapy	Targeted; Chemotherapy
Br38	36.12	36.16	90	50	50	0	ID	2	no	0-1	none	none
Br39	47.38	53.28	60	30	10	0	ID	3	yes	2	Chemotherapy	Hormone; Chemotherapy

Br40	79.82	79.87	100	100	5	0	ID	1	yes	3	Hormone	Hormone
Br41	49.9	61.54	100	100	NA	0	ID	2	yes	1	Hormone	Hormone
Br42	78.76	78.83	80	10	20	0	ID	2	yes	01.Feb	Hormone	Hormone
Br43	69.87	74.73	100	70	25	0	ID	2	yes	2	none	none
Br44	46.84	50.22	100	80	30	1	ID	2	yes	0-1	none	none
Br45	44.68	49.06	100	80	10	0	ID	1	yes	1	none	none
Br46	49.27	49.29	0	0	80	0	ID	3	no	0	none	none
Br47	74.53	77.16	0	0	10	0	ID	3	yes	02.Mar	Targeted; Chemotherapy	Targeted; Chemotherapy
Br48	60.93	65.38	100	100	10	0	ID	1	yes	0	none	none
Br49	77.33	77.37	100	10	20	0	ID	3	no	0	none	none
Br50	52.19	52.22	90	90	5	0	ID	1	no	NA	none	none
Br51	54.13	55.65	0	0	27.5	NA	ID	3	yes	NA	none	none
Br52	48.66	68.15	100	50	35	0	ID	3	yes	1	none	none
Br53	48.22	58.81	1	0	NA	0	ID	NA	yes	1	Hormone	Hormone
Br54	56.91	58.06	100	50	20	0	ID inflammatory	2	yes	1	Chemotherapy	Hormone; Chemotherapy
Br55	62.56	62.63	0	0	15	1	ID	2	yes	0	none	none
Br56	70.09	73.19	50	0	35	0	ID	3	no	3	Chemotherapy	Chemotherapy
Br57	55.39	55.41	95	5	NA	0	ID	3	yes	3	none	none
Br58	66.44	66.48	90	50	30	0	IL	3	yes	0	none	none
Br59	46.32	49.95	100	5	50	1	ID	2	yes	0	Targeted	Targeted; Chemotherapy
Br60	53.88	72.05	100	50	NA	0	NA	NA	yes	1	Hormone	Hormone
Br61	59.82	62.39	60	0	NA	0	ID	2	yes	0-1		Chemotherapy
Br62	80.44	85.61	100	1	20	0	ID	2	yes	01.Feb	Hormone	Hormone
Br63	57.92	60.85	0	0	90	0	ID	3	no	0	Chemotherapy	Chemotherapy
Br64	40.33	50.62	90	2	NA	1	ID	3	yes	1	none	none
Br65	70.02	70.02	0	0	20	0	inflammatory	3	yes	1	none	none
Br66	57.76	74.59	1	NA	NA	0	ID	3	yes	1	Targeted; Hormone	Hormone; Targeted
Br67	56.72	69.65	90	90	NA	0	ID	1	yes	0	Hormone; Targeted	Hormone; Targeted
Br68	46.04	54.13	95	90	NA	0	ID	2	yes	0	none	none
Br69	45.78	45.84	0	NA	10	1	inflammatory	2	yes	0	none	none
Br70	44.61	46.13	100	5	NA	0	ID	2	yes	1	Chemotherapy	Chemotherapy ; Hormone
Br71	61.18	61.16	90	80	30	0	ID	2	yes	NA	none	none
Br72	41.46	42.95	10	0	15	1	ID	3	no	0-1	Chemotherapy; Targeted	Chemotherapy ; Targeted
Br73	79.43	79.49	0	2	20	1	ID	3	yes	0	Targeted; Chemotherapy	Targeted; Chemotherapy
Br74	65.36	68.02	0	70	50	0	ID	2	yes	0-1	Hormone	Hormone
Br75	55.48	60.72	20	0	25	0	ID	2	yes	NA	none	none
Br76	55.4	57.02	0	0	30	0	ID	3	yes	1	Targeted; Chemotherapy	Chemotherapy ; Targeted
Br77	54.63	55.75	100	20	40	0	ID	3	yes	0-1	Hormone	Hormone

Extended Data Table 3. Patient characteristics. The table shows the characteristics of the 73 patients included in the study. *Abbreviations: ER* Estrogen receptor, *HER2* Human epidermal growth factor receptor 2, *ID* Invasive ductal, *IL* Invasive lobular, *NA* Not available, *PR* Progesterone receptor, *ECOG* Eastern Cooperative Oncology Group (as defined by Oken et al. [106]v).

	No CTC (n = 39)	CTC (n = 34)	P	Estimate (95%CI)
Bisphosphonates (%)	9 (23.68%)	7 (20.59%)	0.784	0.84 (0.23 - 2.94)
Denosumab (%)	16 (41.03%)	5 (14.71%)	0.019	0.25 (0.06 - 0.86)
Radiotherapy (%)	22 (56.41%)	13 (38.24%)	0.16	0.48 (0.17 - 1.35)
Therapy at CTC evaluation				
Targeted therapy (%)	8 (20.51%)	6 (17.65%)	1	0.83 (0.21 - 3.13)
Chemotherapy (%)	9 (23.08%)	9 (26.47%)	0.79	1.2 (0.36 - 4)
Hormone therapy (%)	15 (38.46%)	9 (26.47%)	0.324	0.58 (0.19 - 1.74)
Cumulative therapy at CTC evaluation				
Targeted therapy (%)	9 (23.08%)	7 (20.59%)	1	0.87 (0.24 - 3.03)
Chemotherapy (%)	13 (33.33%)	14 (41.18%)	0.628	1.39 (0.48 - 4.05)
Hormone therapy (%)	19 (48.72%)	12 (35.29%)	0.343	0.58 (0.2 - 1.63)

Extended Data Table 4. Therapy evaluation in patients with circulating tumor cells. The table shows the types of therapy that patients with and without circulating tumor cells (CTCs) underwent.

Patient	Bisphosphonates	Denosumab	None
Br1	no	no	yes
Br2	yes	no	no
Br3	no	no	yes
Br4	yes	no	no
Br6	no	no	yes
Br7	no	yes	no
Br8	no	no	yes
Br9	no	no	yes
Br10	yes	yes	no
Br11	no	no	yes
Br12	yes	no	no
Br13	yes	no	no
Br14	yes	yes	no
Br15	no	no	yes
Br16	no	no	yes
Br17	no	no	yes
Br18	no	no	yes
Br19	no	yes	no
Br20	no	no	yes
Br21	no	yes	no
Br22	no	no	yes
Br23	no	no	yes
Br25	no	no	yes
Br27	no	no	yes
Br29	no	no	yes
Br30	no	no	yes
Br31	no	yes	no
Br32	yes	yes	no
Br33	yes	no	no
Br34	no	yes	no
Br35	yes	no	no
Br36	no	no	yes
Br37	no	no	yes
Br38	no	no	yes
Br39	yes	yes	no
Br40	no	no	yes
Br40	no	no	yes
Br41	no	no	yes
Br42	no	no	yes
Br43	no	no	yes
Br44	no	no	yes
Br45	no	no	yes
Br46	no	no	yes
Br47	yes	yes	no
Br48	no	yes	no
Br49	NA	no	NA
Br50	no	no	yes
Br51	no	no	yes
Br52	no	yes	no
Br53	no	no	yes
Br54	yes	yes	no
Br55	no	no	yes
Br56	no	no	yes
Br57	no	no	yes
Br58	yes	yes	no
Br59	yes	yes	no
Br60	no	yes	no
Br61	yes	no	no
Br62	no	no	yes
Br63	yes	no	no
Br64	no	no	yes
Br65	no	no	yes
Br66	no	yes	no
Br67	no	no	yes
Br68	no	no	yes
Br69	no	no	yes
Br70	no	yes	no
Br71	no	yes	no
Br72	no	no	yes
Br73	no	no	yes
Br74	no	yes	no
Br75	no	no	yes
Br76	no	no	yes
Br77	no	yes	no

Extended Data Table 5. Bisphosphonates or denosumab treatment. The table shows whether bisphosphonates or denosumab was administered to each of the patients included in the study.

	Control (n = 52)	Treated (n = 21)	P	Estimate (95%CI)
Age at primary diagnosis, mean (SD)	56.79 (11.7)	57 (11.33)	0.918	-0.31 (-6.34 - 6.01)
Age at first CTC evaluation, mean (SD)	61.12 (11.41)	63.1 (11.3)	0.411	-2.8 (-9 - 3.92)
Stage at Diagnosis			0.213	-
I (%)	7 (13.46%)	2 (10%)		
IA (%)	0 (0%)	1 (5%)		
II (%)	7 (13.46%)	2 (10%)		
IIA (%)	4 (7.69%)	1 (5%)		
III (%)	16 (30.77%)	2 (10%)		
IIIA (%)	0 (0%)	2 (10%)		
IIIC (%)	2 (3.85%)	0 (0%)		
IV (%)	15 (28.85%)	10 (50%)		
Lymphocyte node involvement			0.065	-
N0 (%)	14 (29.17%)	5 (27.78%)		
N1 (%)	15 (31.25%)	10 (55.56%)		
N2 (%)	4 (8.33%)	2 (11.11%)		
N3 (%)	14 (29.17%)	1 (5.56%)		
Histologic subtype			0.077	-
Invasive lobular (%)	7 (13.46%)	3 (14.29%)		
Invasive ductal (%)	43 (82.69%)	17 (80.95%)		
Inflammatory invasive lobular (%)	0 (0%)	2 (9.52%)		
Inflammatory (%)	2 (3.85%)	2 (9.52%)		
% of ER+ cells, mean (SD)	59.19 (41.11)	72.48 (42.25)	0.06	-10 (-20 - 0)
% of PR+ cells, mean (SD)	30.43 (37.99)	39.3 (36.19)	0.168	-5 (-29 - 0)
% of KI67+ cells, mean (SD)	30 (21.2)	27.31 (13.79)	0.941	0 (-10 - 10)
HER2+ (%)	11 (22.45%)	3 (14.29%)	0.52917	0.58 (0.09 - 2.58)
Triple - (%)	6 (12.5%)	1 (5%)	0.66441	0.37 (0.01 - 3.42)
Tumor grade			0.704	-
1 (%)	6 (11.76%)	2 (10%)		
2 (%)	21 (41.18%)	10 (50%)		
3 (%)	24 (47.06%)	8 (40%)		
Bisphosphonates (%)	8 (15.69%)	8 (38.1%)	0.05952	3.24 (0.87 - 12.27)
Radiotherapy (%)	24 (46.15%)	11 (52.38%)	0.79635	1.28 (0.41 - 4.02)
Relapse				
Any (%)	40 (76.92%)	16 (76.19%)	1	0.96 (0.26 - 4.06)
Local (%)	4 (7.69%)	3 (14.29%)	0.40264	1.98 (0.26 - 12.99)
Metastasis (%)	33 (63.46%)	12 (57.14%)	0.79082	0.77 (0.24 - 2.49)
Days between primary diagnosis and relapse, mean (SD)	1560.57 (1491.24)	2995.75 (2680.45)	0.213	-1017.31 (-3417 - 359)
Established metastatic disease at CTC evaluation (%)	44 (84.62%)	21 (100%)	0.09526	Inf (0.73 - Inf)
Number of metastatic sites, mean (SD)	1.96 (0.9)	2.1 (1.14)	0.777	0 (-1 - 0)
Metastasis site				
Bone (%)	24 (47.06%)	20 (95.24%)	5.60E-05	22.53 (3.14 - 995.64)
Liver (%)	14 (27.45%)	8 (38.1%)	0.40326	1.66 (0.49 - 5.49)
Lymphnode (%)	12 (23.53%)	7 (33.33%)	0.38892	1.65 (0.46 - 5.72)
Pleural (%)	7 (13.73%)	2 (9.52%)	1	0.68 (0.06 - 4.03)
Peritoneal (%)	7 (13.73%)	0 (0%)	0.18165	0 (0 - 1.65)
Lung (%)	3 (5.88%)	5 (23.81%)	0.03927	4.97 (0.86 - 35.61)
Skin (%)	3 (5.88%)	0 (0%)	0.55229	0 (0 - 6.05)
Brain (%)	3 (5.88%)	1 (4.76%)	1	0.82 (0.01 - 10.92)
Uterus (%)	2 (3.92%)	0 (0%)	1	0 (0 - 13.31)
Muscular (%)	3 (5.88%)	0 (0%)	0.55229	0 (0 - 6.05)

Extended Data Table 6. Clinical features of patients in regard to denosumab treatment. The table shows clinical features of patients who were treated or not with denosumab. *Abbreviations: ER* Estrogen receptor, *HER2* Human epidermal growth factor receptor 2, *PR* Progesterone receptor.

	CTC single cell (n = 22)	CTC clusters (n = 12)	P	Estimate (95%CI)
Age at primary diagnosis, mean (SD)	57.54 (9.52)	50.63 (12.6)	0.04	-8.6 (-16.08 - -0.31)
Age at first CTC evaluation, mean (SD)	62.15 (9.11)	54.87 (12.14)	0.06	-7.73 (-15.37 - 0.42)
Stage at Diagnosis			0.85	-
I (%)	3 (13.64%)	2 (16.67%)		
IA (%)	0 (0%)	0 (0%)		
II (%)	2 (9.09%)	2 (16.67%)		
IIA (%)	3 (13.64%)	1 (8.33%)		
III (%)	5 (22.73%)	2 (16.67%)		
IIIA (%)	0 (0%)	0 (0%)		
IIIC (%)	1 (4.55%)	1 (8.33%)		
IV (%)	7 (31.82%)	4 (33.33%)		
Lymphocyte node involvement			0.9	-
N0 (%)	4 (20%)	4 (36.36%)		
N1 (%)	11 (55%)	3 (27.27%)		
N2 (%)	0 (0%)	0 (0%)		
N3 (%)	5 (25%)	4 (36.36%)		
Histologic subtype			0.7	-
Invasive lobular (%)	3 (13.64%)	1 (8.33%)		
Invasive ductal (%)	18 (81.82%)	11 (91.67%)		
Inflammatory invasive lobular (%)	0 (0%)	1 (8.33%)		
Inflammatory (%)	1 (4.55%)	1 (8.33%)		
% of ER+ cells, mean (SD)	56.18 (40.82)	66.42 (42.1)	0.36	1 (-10 - 30)
% of PR+ cells, mean (SD)	22.18 (31.81)	34 (37.3)	0.36	1 (-4 - 40)
% of KI67+ cells, mean (SD)	34.12 (22.64)	23 (16.43)	0.21	-10 (-25 - 10)
HER2+ (%)	6 (30%)	1 (9.09%)	0.37	0.24 (0 - 2.52)
Triple - (%)	3 (15%)	0 (0%)	0.54	0 (0 - 4.4)
Tumor grade			0.07	-
1 (%)	1 (4.55%)	3 (27.27%)		
2 (%)	9 (40.91%)	5 (45.45%)		
3 (%)	12 (54.55%)	3 (27.27%)		
Bisphosphonates (%)	5 (22.73%)	2 (16.67%)	1	0.69 (0.06 - 5.25)
Denosumab (%)	3 (13.64%)	2 (16.67%)	1	1.26 (0.09 - 12.99)
Radiotherapy (%)	8 (36.36%)	5 (41.67%)	1	1.24 (0.23 - 6.52)
Relapse				
Any (%)	16 (72.73%)	9 (75%)	1	1.12 (0.18 - 8.63)
Local (%)	1 (4.55%)	3 (25%)	0.12	6.57 (0.46 - 383.48)
Metastasis (%)	14 (63.64%)	5 (41.67%)	0.29	0.42 (0.08 - 2.14)
Days between primary diagnosis and relapse, mean (SD)	1996.2 (2006.19)	1636.67 (1538.48)	0.62	-291.5 (-1975 - 1591)
Established metastatic disease at CTC evaluation (%)	19 (86.36%)	11 (91.67%)	1	1.71 (0.12 - 99.29)
Number of metastatic sites, mean (SD)	1.75 (0.91)	2.18 (0.98)	0.21	0 (0 - 1)
Metastasis site				
Bone (%)	10 (47.62%)	7 (58.33%)	0.72	1.65 (0.33 - 8.94)
Liver (%)	9 (42.86%)	3 (25%)	0.47	0.49 (0.07 - 2.77)
Lymphnode (%)	6 (28.57%)	4 (33.33%)	0.71	1.32 (0.21 - 7.67)
Pleural (%)	2 (9.52%)	0 (0%)	0.53	0 (0 - 9.85)
Peritoneal (%)	2 (9.52%)	2 (16.67%)	0.6	1.96 (0.12 - 30.79)
Lung (%)	3 (14.29%)	1 (8.33%)	1	0.58 (0.01 - 8.36)
Skin (%)	0 (0%)	0 (0%)	1	0.58 (0.01 - 8.36)
Brain (%)	1 (4.76%)	1 (8.33%)	1	1.87 (0.02 - 156.66)
Uterus (%)	0 (0%)	1 (8.33%)	0.35	Inf (0.05 - Inf)
Muscular (%)	1 (4.76%)	1 (8.33%)	1	1.87 (0.02 - 156.66)

Extended Data Table 7. Clinical features of patients with single circulating tumor cell and circulating tumor cell clusters. The table shows clinical features of patients in whom only single circulating tumor cells (CTC single cell) or also clustered circulating tumor cells (CTC clusters) were found. *Abbreviations: ER* Estrogen receptor, *HER2* Human epidermal growth factor receptor 2, *PR* Progesterone receptor.

	Single CTC (n = 22)	CTC clusters (n = 12)	P	Estimate (95%CI)
CA 15-3, mean (SD)	66.15 (57.74)	2554.6 (6387.64)	0.0089	234.2 (17.4 - 540.4)
Alkaline phosphatase, mean (SD)	108.47 (108.56)	310.25 (525.57)	0.4941	9 (-18 - 109)
Calcium (korr), mean (SD)	2.29 (0.11)	2.36 (0.35)	0.8835	0.01 (-0.14 - 0.21)
CRP, mean (SD)	17.96 (35.81)	37.25 (58.6)	0.1497	4.99 (-2 - 31.5)
LDH, mean (SD)	246 (43.64)	363.33 (329.24)	0.3035	30.66 (-20 - 96)
RBC 10^12/L, mean (SD)	3.99 (0.63)	3.66 (0.93)	0.4996	-0.22 (-1.08 - 0.31)
HGB g/L, mean (SD)	121.07 (20.19)	114.18 (29.19)	0.6777	-5 (-29 - 14)
HCT L/L, mean (SD)	0.36 (0.04)	0.34 (0.08)	0.7204	-0.02 (-0.09 - 0.03)
MCV fL, mean (SD)	88.27 (5.48)	91.73 (4.41)	0.081	4 (-2 - 8)
MCH pg, mean (SD)	30.23 (2.19)	31.24 (1.51)	0.1608	1 (-0.6 - 2.6)
MCHC g/L, mean (SD)	341.6 (12.12)	340.64 (14.25)	0.8557	1 (-13 - 12)
WBC 10^9/L, mean (SD)	5.68 (2.15)	9.38 (3.99)	0.0045	3.48 (1.15 - 6.42)
Neutrophile 10^9/L, mean (SD)	4.03 (1.57)	6.65 (3.6)	0.0306	2.02 (0.11 - 4.9)
Lymphocyte10^9/L, mean (SD)	1.4 (0.76)	1.42 (1)	0.8859	-0.07 (-0.76 - 0.9)
Monocyte 10^9/L, mean (SD)	0.4 (0.15)	0.5 (0.24)	0.2918	0.07 (-0.08 - 0.32)
Eosinophil 10^9/L, mean (SD)	0.17 (0.1)	0.17 (0.14)	0.7844	-0.02 (-0.13 - 0.11)
Basophil 10^9/L, mean (SD)	0.04 (0.03)	0.04 (0.04)	0.6197	-0.01 (-0.03 - 0.02)
LUC 10^9/L, mean (SD)	0.11 (0.04)	0.12 (0.11)	0.4379	-0.02 (-0.06 - 0.02)
PLT 10^9/L, mean (SD)	248.67 (77.76)	251.5 (109.99)	0.9809	2.5 (-77 - 73)
MPV fL, mean (SD)	8.59 (1.89)	8.82 (0.87)	0.153	1 (0 - 1)

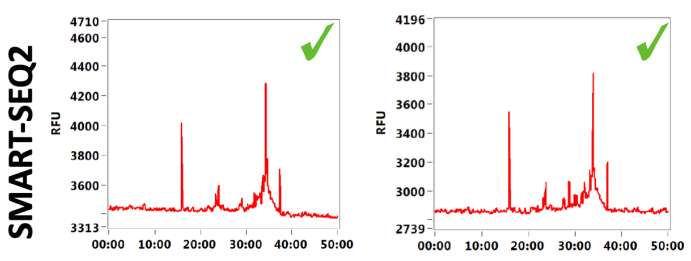
Extended Data Table 8. Complete blood counts in patients with single circulating tumor cells and circulating tumor cell clusters. The table shows complete blood counts in patients in whom only single circulating tumor cells (single CTC) or also clustered circulating tumor cells (CTC clusters) were found. *Abbreviations: CA 15-3* Cancer antigen 15-3, *CRP* C-reactive protein, *HCT* Hematocrit, *HGB* Hemoglobin, *LDH* Lactate dehydrogenase, *LUC* Large unstained cells, *MCH* Mean corpuscular hemoglobin, *MCHC* Mean corpuscular hemoglobin concentration, *MCV* Mean corpuscular volume, *MPV* Mean platelet volume, *PLT* Platelets, *RBC* Red blood cells, *WBC* White blood cells.

5.2. UNPUBLISHED RESULTS

5.2.1 SINGLE CELL RNA AND DNA SEQUENCING OF INDIVIDUAL CTCS

In order to apply parallel DNA and RNA isolation and amplification from single cells, several protocols were tested. The base for method validations was parallel Genome and Transcriptome sequencing (G&T-Seq) [65], (which combined MDA (Multiple Displacement Amplification) for the genome and SMART-Seq2 method for the transcriptome processing. In order to proceed with the most efficient sample preparation, the published protocol was compared side-to-side with two newly introduced commercial methods – AMPLI1 WTA (whole transcriptome amplification) method and MALBAC (Multiple Annealing and Looping Based Amplification Cycles) for the genome amplification. Of note, two whole genome amplification (WGA) kits from the same manufacturer were tested (V3 vs single cell-specific). Interestingly, V3 yielded better results (higher concentrations) for individual cells and was used in further steps.

Based on the final sample concentration and amplified fragments sizes, the original G&T-Seq protocol, including few minor modifications, was chosen for further sample preparation. RNA content was amplified at least 1'000-fold (from around 2 pg to 10-40 ng). Moreover, as opposed to the alternative methods, improved MDA protocol provided over 1ug total gDNA (from starting 6.6 pg), amount necessary for the library construction, and yielded consistent fragment sizes of amplified transcriptome (Figure 1).



MDA

sample	*total ng*
Single cell 1	>3000
Single cell 2	>3000
Single cell 3	>3000
Single cell 4	>3000
15 cells 1	>3000
15 cells 2	>3000
NEG ctrl	0

MALBAC

sample	*total ng*
Single cell 1	493
Single cell 2	430
Single cell 3	493
Single cell 4	383
15 cells 1	595
15 cells 2	590
NEG ctrl	7

Figure 1. Quality control of single-cell WGA. Representative distribution of fragment sizes for SMART-Seq2 and AMPLI1 (left). Concentrations of gDNA obtained after the amplification with either SMART-Seq2 and

MALBAC approaches (right). Green check mark indicates results that were necessary for further protocol steps. Red crosses point towards poor quality data.

Initial exome and transcriptome (E&T) validation sequencing with HiSeq 2500 was performed on CTC-derided cell line (Brx68-GFP) and resulted in high quality data (Figure 2 and Table 1).

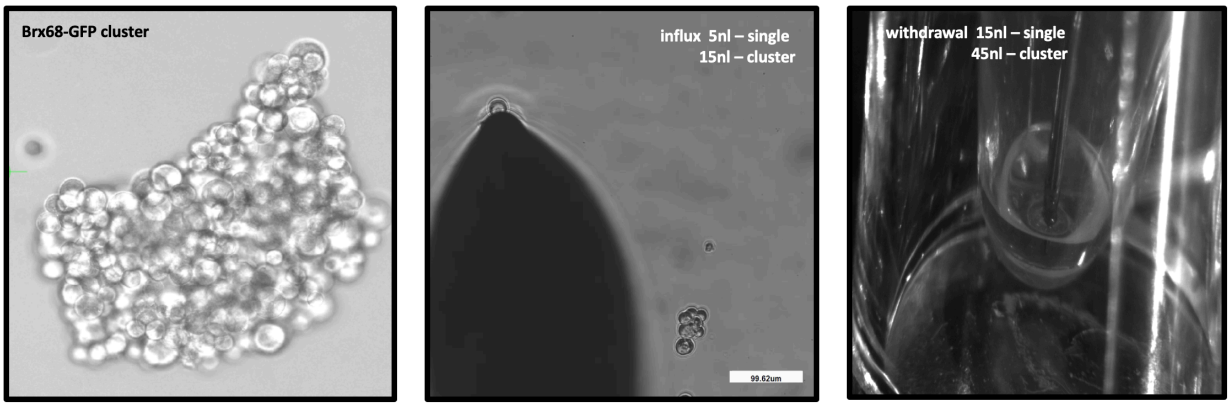


Figure 2. First attempts in isolation of individual cells. Pictures captured during the manipulation of cells used for E&T-seq protocol validation. Courtesy of Dr. Stefan Arnold.

There were on average 9832 transcripts detected from 7 different single cells, with a minimum of 10 reads per each transcript (depth). In addition, protocol included two positive controls of 15-cell pools (to ensure the presence of deposited cell), and RNA sequencing of these samples resulted in the mean of 13333 transcripts (10x depth).

Sample ID	*Genes With 1 Read*	*Genes With 10 reads*	*Genes With 20 reads*
15 cells 1	15611	13536	12950
15 cells 2	15718	13130	12570
Single cell 1	11943	10554	10205
Single cell 2	10564	9052	8803
Single cell 3	10782	9350	9001
Single cell 4	11399	9974	9610
Single cell 5	12153	10782	10429
Single cell 6	10336	9098	8821
Single cell 7	11356	10019	9726

Table 1. E&T-Seq validation. Number of genes detected with 1,10 or 20 reads are shown for 15-cell pools (in duplicate) and 6 single cells. *Note:* Computational analysis performed by Dr. Robert Ivanek.

The quality of amplified transcriptome has been assessed with AMPLI1 quality control (QC) kit (Figure 3).

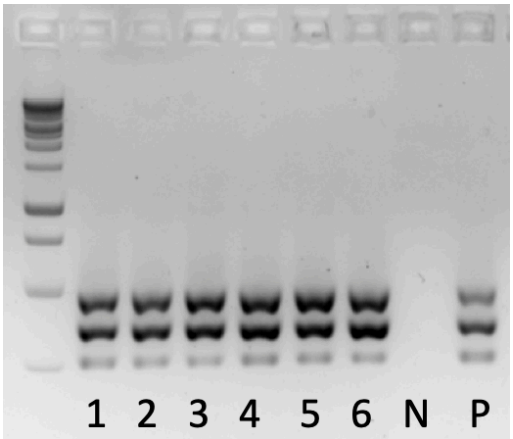


Figure 3. Whole transcriptome amplification QC. An agarose gel visualizing QC PCR product. Three bands represent good quality amplification products and are recommended for samples submitted to RNA sequencing. Samples 1-6 represent six individual CTCs (mouse-derived), N = negative control (no template), P = positive control (spiked gDNA).

Correlation analysis of transcripts shared between all sequenced samples proved high protocol reproducibility (74-90% identical sequences detected between samples). Importantly, the additional of internal RNA control (ERCC spike-in) enabled more precise sample analysis and the concentration of tested spike-in was approved for all the future samples (Figure 4).

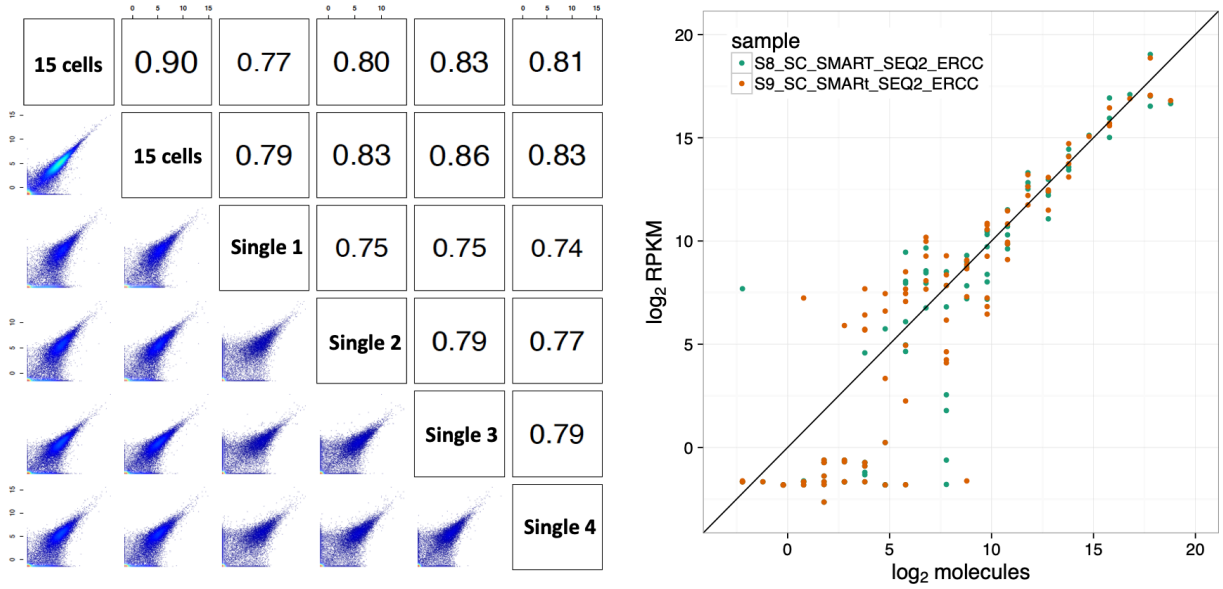


Figure 4. RNA-seq data quality. Correlation plot showing proportion of shared sequences between samples (left). Sensitivity of the detection of RNA expression across different samples is analyzed by evaluating the expression of both ERCC spike-in control transcripts (right). *Note:* Computational analysis performed by Dr. Robert Ivanek.

On the other hand, whole exome sequencing from matched samples resulted in 33-63% of exome coverage for single cells and 80-85% for 15-cell pools at 5x depth (Figure 5).

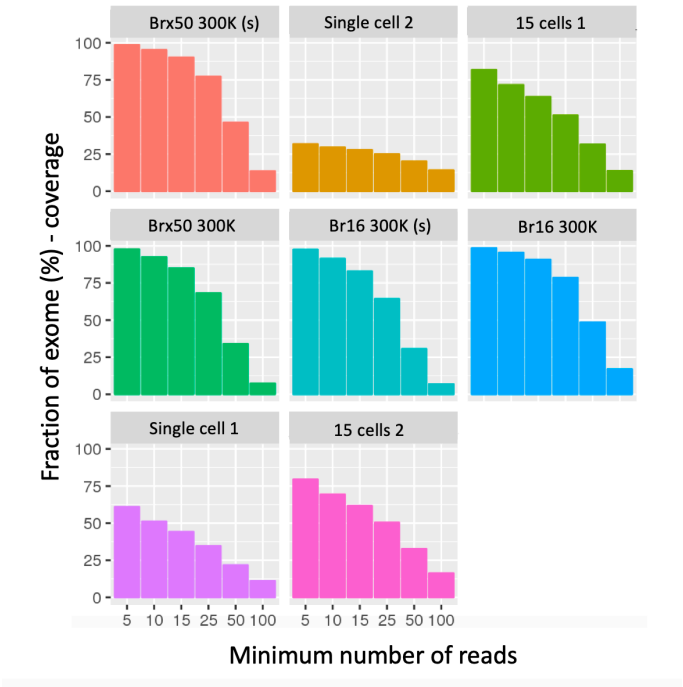


Figure 5. DNA Data quality. Fraction of exome coverage at different number of reads shown for multiple samples ranging between single cell and 300'000 cells. *Note:* Computational analysis performed by Dr. Robert Ivanek.

Interestingly, mouse (xenograft)-derived CTCs provided better quality DNA-seq data compared to patient CTCs, as evaluated based on the % coverage of the exome. Taken together, these data confirm a proper protocol execution for E&T-Seq (Figure 6).

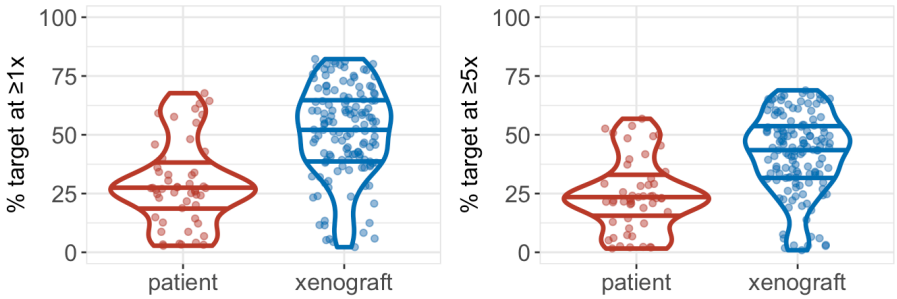
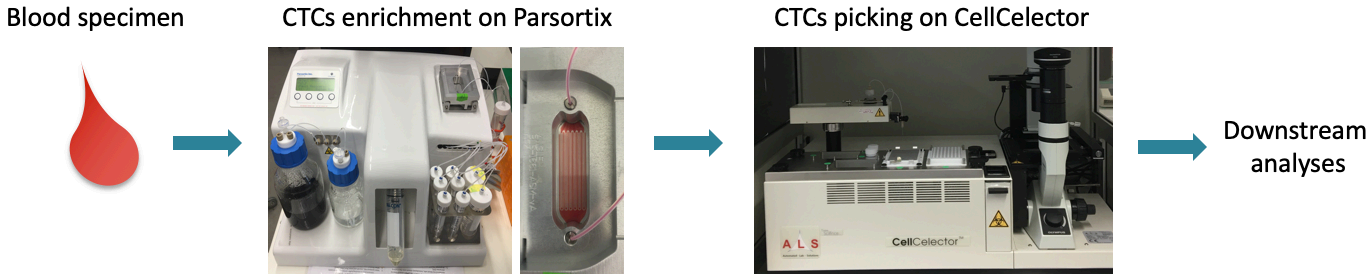


Figure 6. Mouse-derived CTCs provide better quality DNA-Seq data. Boxplots showing the percentage of exome coverage distribution between patient- and xenograft (mouse)– derived CTCs, at either 1x (left) or 5x depth. *Note:* Computational analysis performed by Dr. Francesc Castro-Giner.

5.2.2 PATIENT AND MOUSE-DERIVED CTC ISOLATION

Circulating tumor cells were isolated from freshly drawn blood of cancer patients and mouse models using two-step procedure.



Firstly, blood was passed through a microfluidic system (Parsortix) in order to enrich large tumor cells and remove the majority of unwanted red and white blood cells. Secondly, collected CTCs were deposited into ultra-low attachment and individual cells of interest were micromanipulated (CellCelector) to allow downstream analyses (Figure 7).

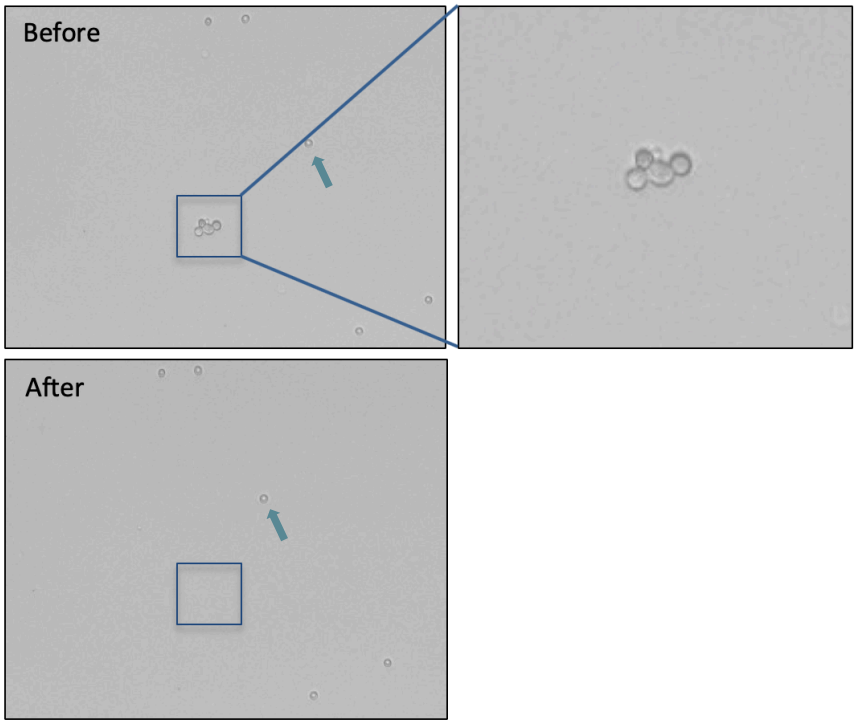


Figure 7. Precision of cell micromanipulation with CellCelector. Representative pictures form before (top) and after (bottom) cell cluster micromanipulation are shown. Red blood cell in proximity of isolated CTC cluster is indicated with an arrow.

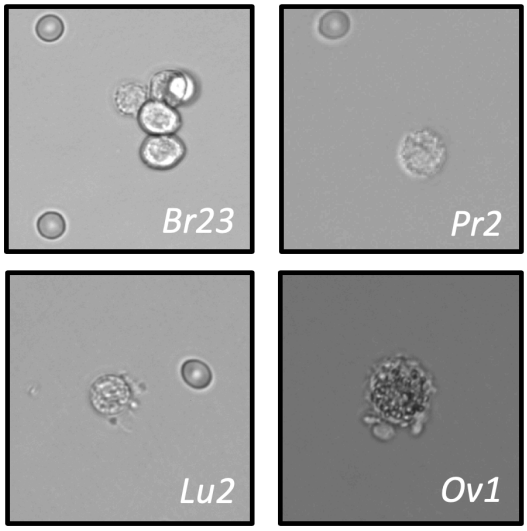


Figure 8. Representative images of patient-derived CTCs. CTC clusters (top left) or single CTCs (top right and both bottom) were imaged before micromanipulation. Cells originated from either breast (Br), prostate (Pr), lung (Lu) or ovarian (Ov) cancers. Patient ID is shown after the cancer type.

Occasionally, cell distress was detected based on the morphology (e.g. in Figure 8, Lu2 and Ov1), probably due to the ongoing chemotherapy. On the other hand, CTCs from drug naive mouse models displayed a healthier phenotype (Figure 9).

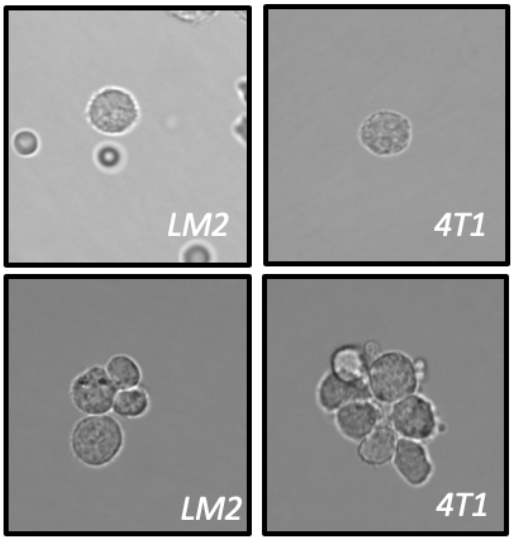


Figure 9. Representative images of mouse-derived CTCs. Single CTCs (top) and CTC clusters were isolated at the time of an overt metastatic disease from NSG mice injected with either LM2 (left) or 4T1 (right) cells.

In case of clustered cells, micromanipulation was used to mechanically dissociate attached cells using aspiration and immediate release into the capillary, performed sequentially around ten times. This procedure did not induce any visible changes in cell morphology (Figure 10).

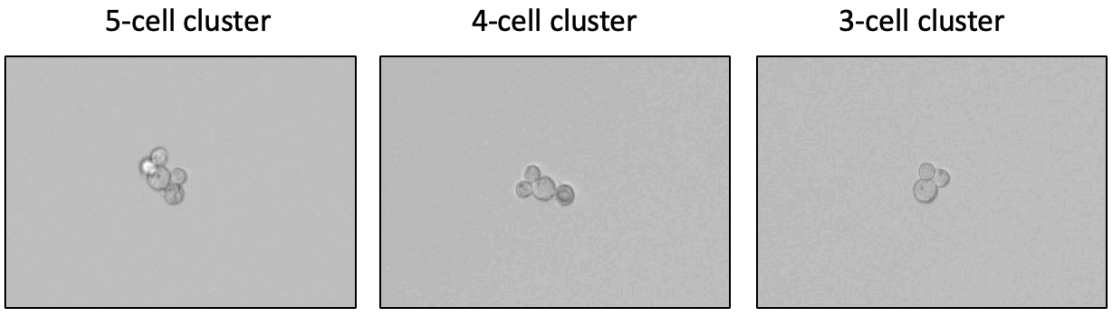


Figure 10. Isolation of cells from CTC clusters. Individual cells were subsequently removed from a 5-cell cluster from NSG-BR16 mouse model. *Note*: 3-cell cluster from this example was unmanageable to dissociate further and was sequenced as a bulk of three cells.

A total number of CTCs collected over the course of the study is shown in Table 2 for patients and Table 3 for mouse models.

	Breast	**Lung**	**Prostate**	**Ovarian**
Donor #	**18**	**2**	**5**	**2**
Single CTCs	156	28	41	5
CTC clusters	52	0	6	2
CTC-WBC cluster	25	0	4	0

Table 2. Summary of CTC collection from cancer patients. Single and clustered CTCs were collected from breast, lung, prostate and ovarian cancer patients. Numbers of donors from each cancer type are shown in bold. A sum of collected single CTCs, CTC clusters and CTC-WBC clusters are given below each patient group.

	NSG-BR16	**NSG-BRx50**	**NSG-LM2**	**NSG-4T1**	**BALB/c-4T1**	**PyMT**
Single CTCs	32	43	51	57	16	1
CTC clusters	40	25	44	58	6	18
CTC-WBC cluster	7	3	24	15	11	6

Table 3. Summary of CTCs collected from mouse models. Table shows total numbers of collected single CTCs, CTC clusters and CTC-WBC clusters for different mouse models used in the study.

CTCs obtained from different mouse models were consistent in respect of cell size. An exception was a transgenic MMTV-PyMT model with spontaneous cancer in multiple mammary fat pads. Isolated cells were found much smaller compared to other models (15-20 µm in diameter for BR16 vs. 7-10 µm for MMTV-PyMT), which perhaps can explain low capture rate in Parsortix system (only one single CTC found among 24 CTC clusters). Interestingly, PyMT-derived clusters formed "tight" multicellular spheres, as opposed to "string-like" clusters of BR16-model (Figure 11).

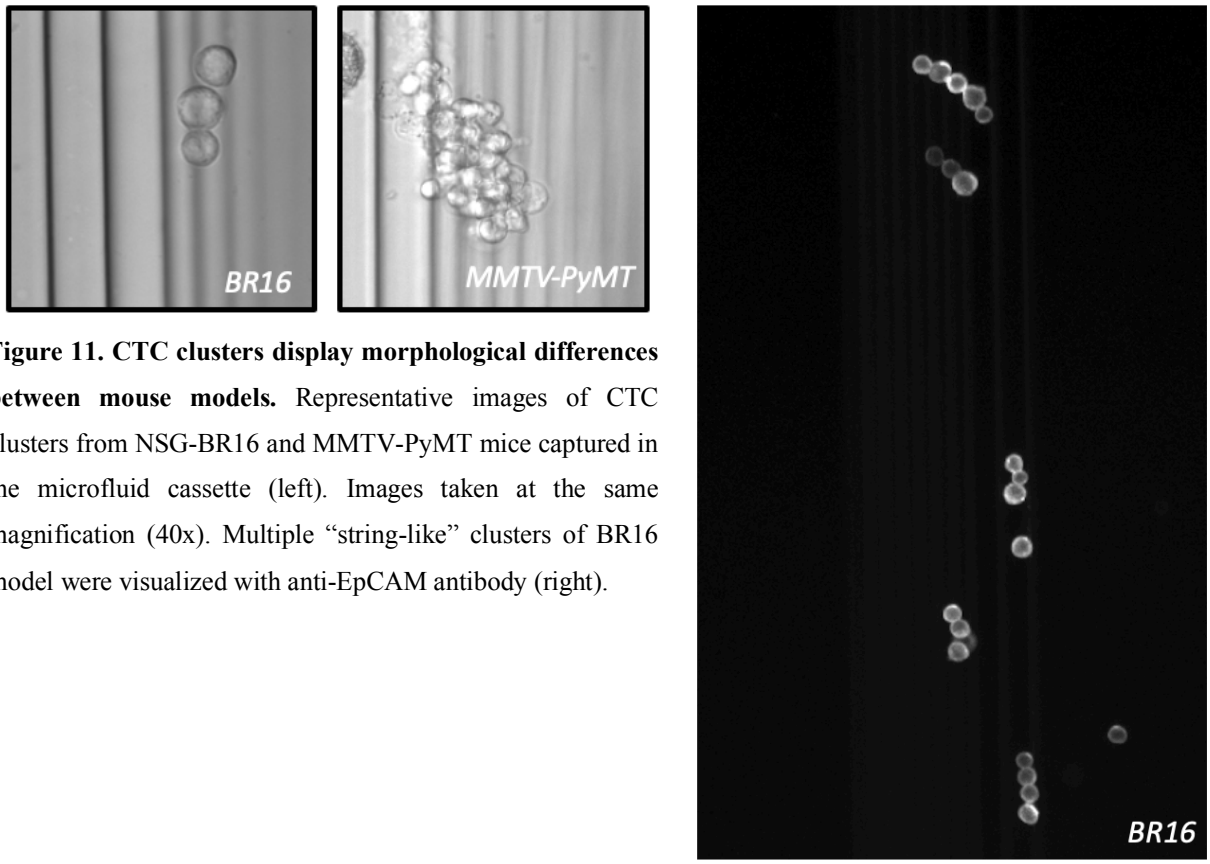


Figure 11. CTC clusters display morphological differences between mouse models. Representative images of CTC clusters from NSG-BR16 and MMTV-PyMT mice captured in the microfluid cassette (left). Images taken at the same magnification (40x). Multiple "string-like" clusters of BR16 model were visualized with anti-EpCAM antibody (right).

5.2.3 SINGLE CELL VS. CELL CLUSTERS PROLIFERATION AND SURVIVAL POTENTIAL

The role of cell-cell connection in survival and proliferation was studied in vitro in CTC-derived cell lines (Brx50, Brx68 and BR16). The growth of over 1000 single cells and cell clusters (2-14 cells) was monitored during 8-week period in 384-well format (Figure 12).

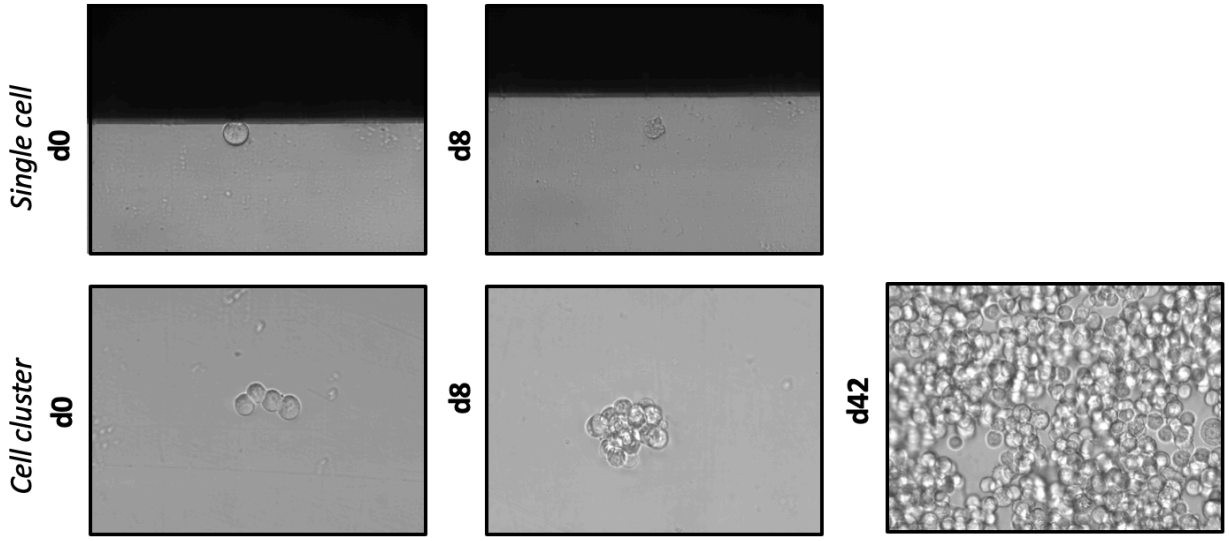


Figure 12. Proliferation of single vs. clustered cells. Representative images of Brx50 cells seeded either individually (top) or as clusters (bottom). Significant improvement in survival and proliferation was observed for clustered cells.

Obtained data was normalized by a total number of cells included in the experiment. This experiment showed clear survival advantage of cell clusters, with 18% (clusters) vs. 4% (single cells) survival rate for Brx50. Similar results were observed for Brx68, approximately 90% of clusters and only 40% of single cell survived over 56 days (Figure 13).

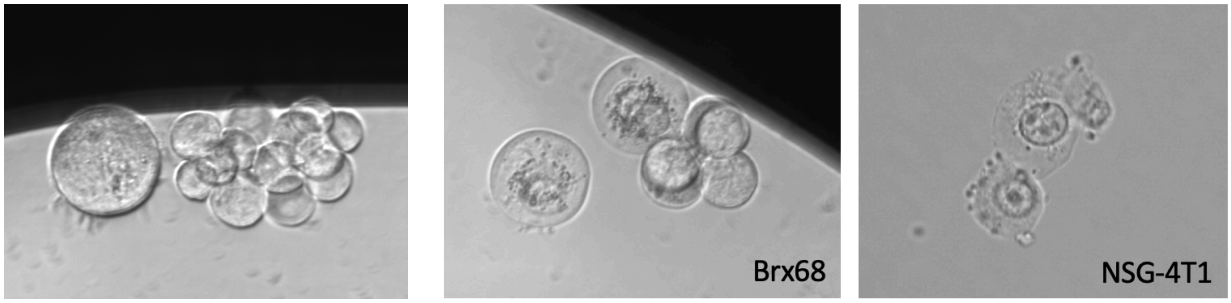


Figure 13. Distinct morphology of Brx68 cells. Single-cell derived colony shows dramatic difference after the first division: dormant "giant" cell vs. proliferating progeny of a sister cell (left). "Egg-like" viable cells present in the culture (middle). Similar structures were found also among NSG-4T1 (right) and NSG-LM2 CTCs.

Rare single-cell derived colonies exhibited much lower total cell numbers when compared to normalized (i.e. final cell number was divided by the number of cells at the time of seeding) colonies originating from cell clusters (Figure 14).

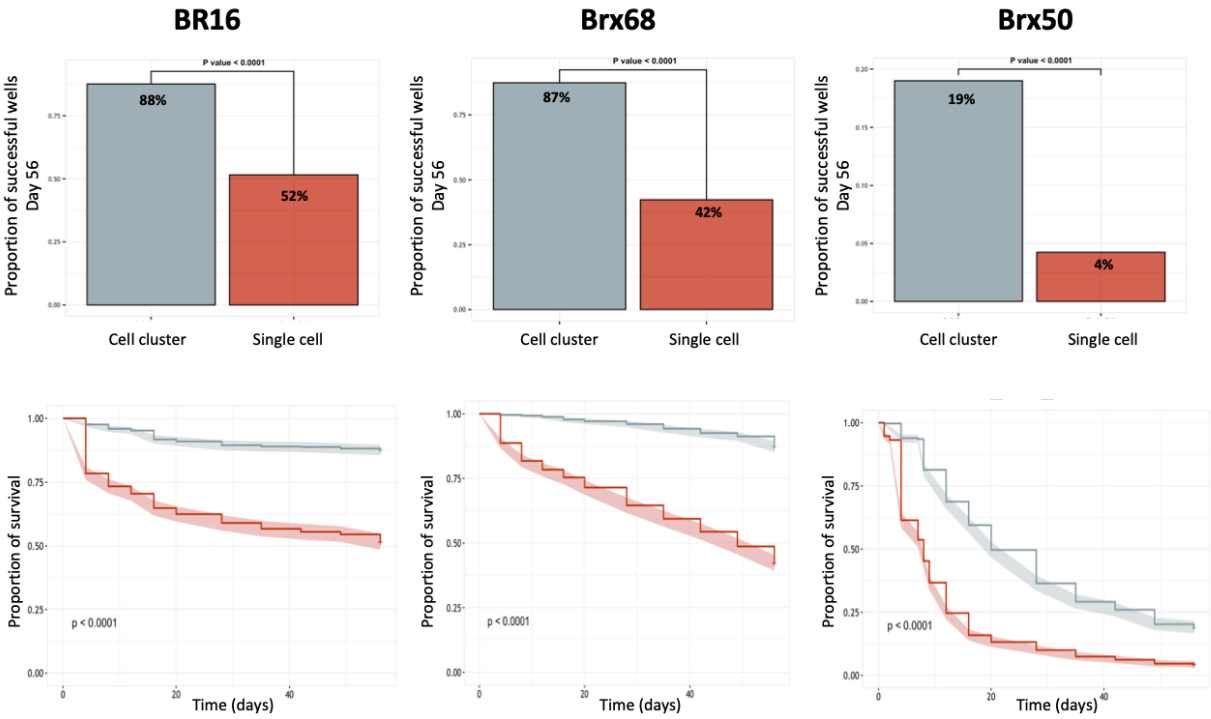


Figure 14. Survival advantage of CTC clusters. Proportions of successful wells (i.e. with live cells at day 56) are shown for the cell lines (BR16, Brx68 and Brx50) (top). The same experiment was summarized with Kaplan-Meier survival analysis (bottom).

Cells with the highest proliferation were used to create clone-derived sublines (26 cluster-derived and one single cell-derived for Brx50 and 24 each for BR16) and tested in mice for metastatic potential (study not yet concluded). Moreover, this experiment enabled identification of quiescent cells defined by no cell division for 56 days. These cells were isolated and preserved for further analysis. Morphological abnormality was observed (Figure 15).

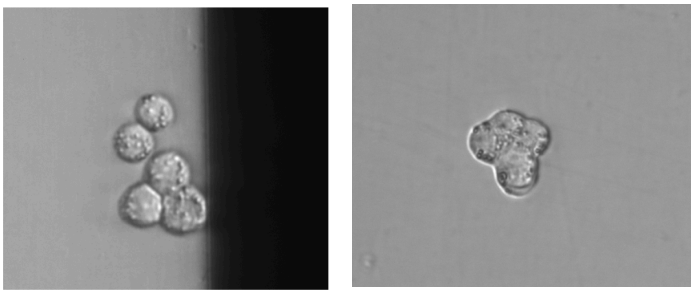


Figure 15. Atypical BR16 cells. Representative pictures showing granular surface (left) and incomplete division (right) found among some proliferating BR16 colonies.

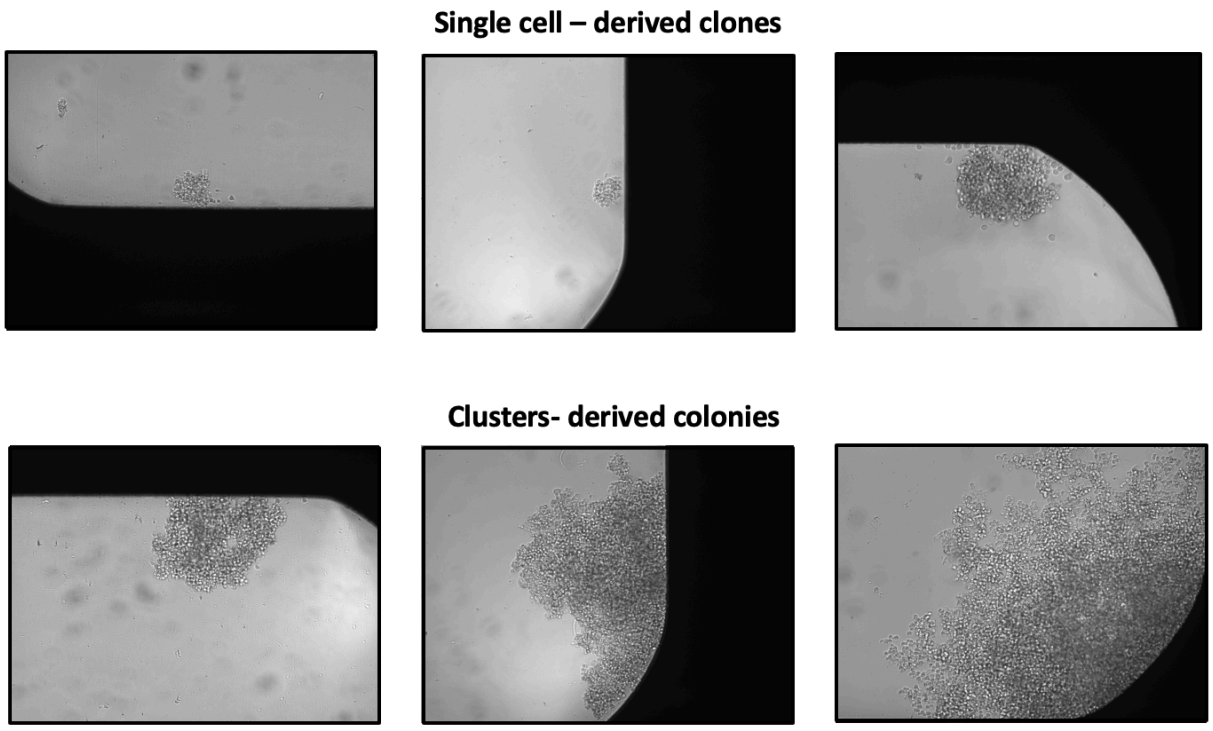


Figure 16. Size difference between single cell- and clustered cell-derived colonies. Representative images of Brx50 colonies.

5.2.3 PROLIFERATION EFFICIENCY OF CTC-DERIVED CELL LINE UNDER DIFFERENT CONCENTRATION OF OXYGEN

Establishment of a patient CTC-derived cell line is known to be challenging due to lack of proliferation or immediate cell death. Even existing protocols have multiple limitations and struggle with low culture efficiency [68, 107], noted also during culturing trials of patient samples from the "CTC-study" (e.g. unsuccessful Br7, Ov10, Lu2, Pr2 cultures). Since hypoxia is a known regulator of metastasis [108], one of the characteristic features in CTC-specific culture is a decrease of oxygen supply to 1-5%. In order to test how CTC-derived cell line (Brx50) will react to changes in oxygen, equal number of cells was cultured at either standard condition (5% oxygen) vs. significant hypoxia (1% oxygen) or normoxia (21% oxygen) shown in Figure 17.

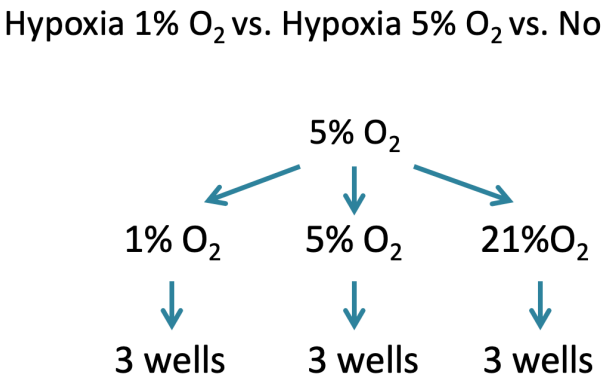


Figure 17. Schematic of experimental design. Brx50 cells were transferred from standard culture conditions into lower oxygen concentration or normoxia. Experient was performed in triplicates.

After 21 days of culture results indicated that previously maintained oxygen concentration was optimal for cell proliferation. Although cells tended to cluster more at 20% oxygen, their total numbers were lower than at previous hypoxic conditions. Interestingly, further limitation of oxygen down to 1% slowed down the cell cycle the most. Cells did not proliferate, but at the same time they did not exhibit any morphological signs of distress over the three-week culture. This could be the result of intrinsic or acquired *in vitro* oxygen-dependency for Brx50 cell line.

Further studies are required to decipher the influence of oxygen levels on physiology of CTC-derived cell lines (Figure 18).

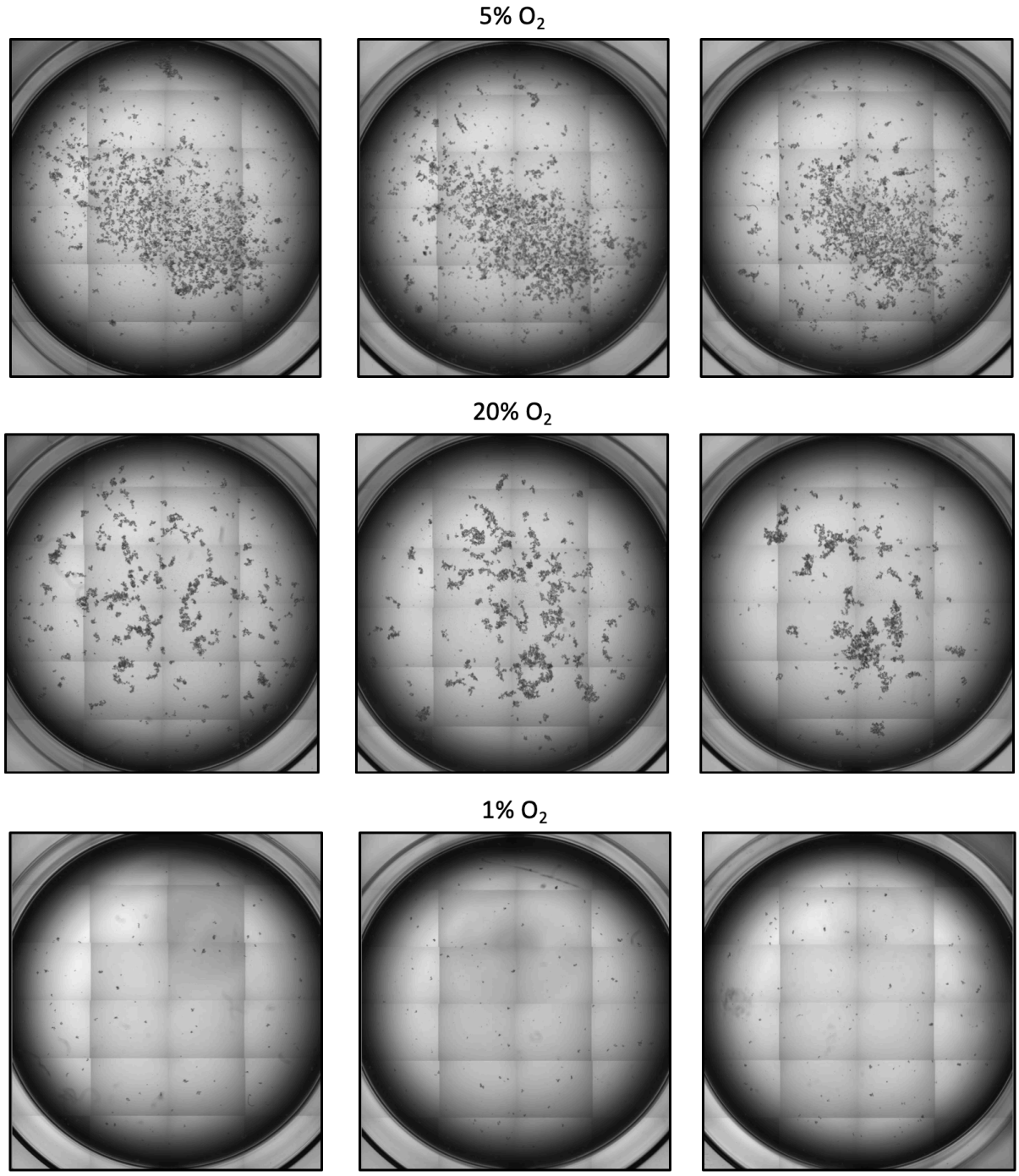


Figure 18. Oxygen dependence for CTC growth. Pictures of Brx50 cells cultured at different oxygen levels were taken after 21 days. All technical replicates are shown (lines). Pictures were taken and stitched using the CellCelector software.

5.2.4 COMPLETE BLOOD COUNTS (CBCS) ACROSS DIFFERENT MOUSE MODELS

Significant changes in the total number of blood cells are known to occur in cancer patients, especially due to the chemotherapy [109]. However, pre-therapy blood counts are known for decades for their prognostic significant in cancer field, e.g. high white blood cell counts corelate with worse survival [110-113].

To address the impact of cancer on CBC changes, the blood was taken from either naïve or tumor-bearing mice, at the time of experiment termination (Figure 19).

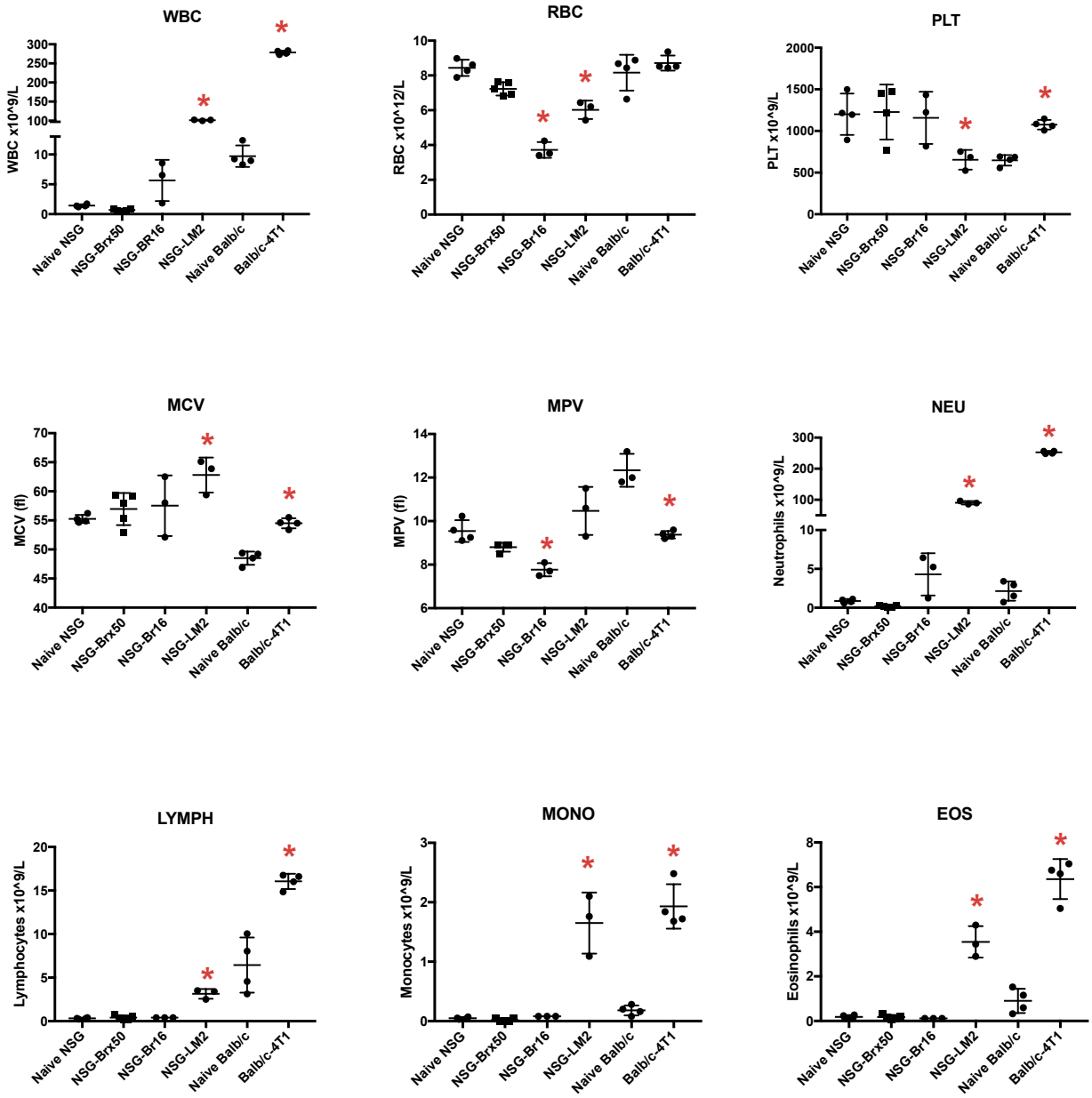


Figure 19. CBCs among breast cancer mouse models. Blood counts were measured for control mouse strains (naïve NSG and BALB/c) or tumor-bearing mice at the time of an overt metastatic disease. WBC – white blood cell, RBC – red blood cell, PLT – platelet, MCV – mean corpuscular volume, MPV – mean platelet volume, NEU – neutrophil, LYMPH – lymphocyte, MONO – monocyte, EOS – eosinophil. Significant differences with naïve mice were marked with red star.

Across mice with NSG background Brx50 model had no significant impact on CBCs. Notably, this CTC-derived cell line although tumorigenic, rarely causes metastatic disease. Similarly, NSG-BR16 mice did not stand out except lower red blood cell numbers and smaller platelet size. Mice injected with LM2 cells exhibited the biggest differences. Namely, higher count of WBCs, seen across all measured leukocytes (neutrophils, monocytes and eosinophils), lower red blood cell and platelet counts and bigger corpuscular volume. Importantly, higher counts of lymphocytes detected upon LM2 tumor growth are the result of an inaccurate cell assignment by the CBC enumeration system, counting most likely abnormal monocytes and granulocytes. In case of BALB/c background only one parameter (number of RBCs) was not changed indicating dramatic changes induces by 4T1 cancer growth in these immunocompetent mice.

These findings can be linked to the changes of spleen size in different models, as spleen is a known immune cell storage organ (Figure 20).

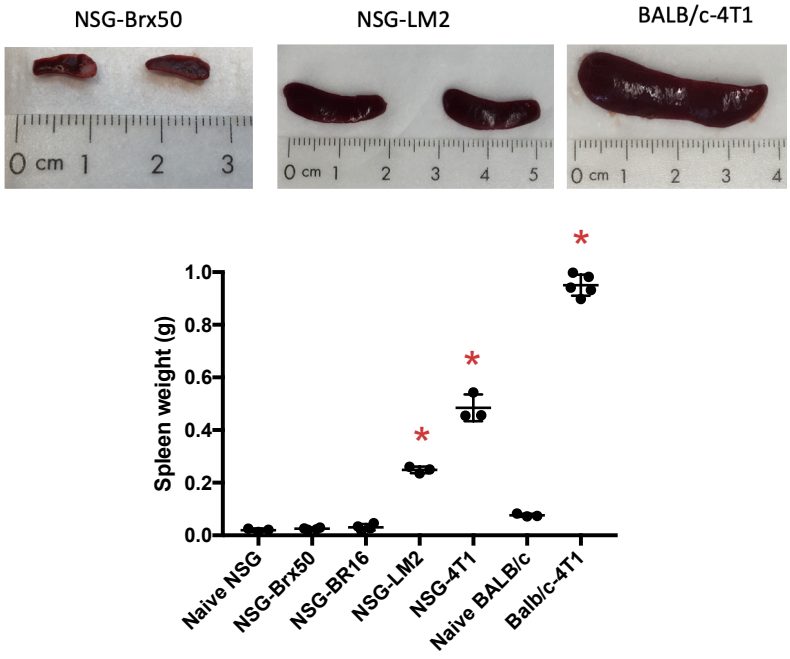


Figure 20. Spleen size across mouse models. Representative pictures of spleens from different breast cancer mouse models (top). Graph showing the weight of spleens for each model (bottom).

6. CONCLUSIONS

The ability to metastasize to distant sites is a fatal hallmark and possibly one of the least understood and most complex features of cancer. Due to the continuous interest and recent development of novel technologies our understanding of metastatic mechanisms is rapidly broadening, especially in the context of CTC biology. Scientists are now offered a long-anticipated opportunity to study these rare cells and tackle the key questions in cancer field.

A unique approach applied in the described study allowed simultaneous analyses of exome and transcriptome at the single-CTC level. This enabled deeper characterization of highly understudied cells leading to the description of pro-metastatic interactions between CTCs and neutrophils. The ability of cancer to take such an advantage of immune cells might explain why certain tumors are more aggressive than other and how CTC vulnerability might be overcome in the circulation. These results are in line with previously reported prognostic value of CTC-WBC clusters, showing that cancer cells can carry the tumor "soil" to facilitate metastasis and worsen prognosis. Moreover, by focusing on cell clusters (as opposed to single CTCs) and highlighting the role of the rarest CTC subtypes, emphasis was placed at the highly relevant cell entities, that were previously disregarded. Notably, interaction with other immune cells were observed, opening yet another area for exploration (CTC-T cell or CTC-B cell clusters).

Findings published in *Nature* are not restricted only to breast cancer. The same workflow can be applied to other metastatic cancers. In fact, multiple samples from lung, prostate and ovarian cancers were already collected during this study. Additionally, provided approach can be applied to assess the exome and transcriptome profile changes along the therapy, leading to identification of new drug resistance mechanisms or prediction of drug responsiveness. In the era of *personalized medicine*, liquid biopsy is bringing hope for non-invasive and precise clinical tool. Importantly, collected parallel DNA and RNA sequencing enables future analysis of the direct impact of mutational profile on gene expression at a single-cell level.

Data provided in the second manuscript points towards the previously unknown association between CTC absence and the treatment with the monoclonal antibody (Denosumab) among breast cancer patients.

This finding suggests that Denosumab treatment may be beneficial to reduce cancer spread in patients that are diagnosed with bone metastasis. Moreover, while factors such as limited blood volume and diverse CTC-isolation technologies may influence CTC detection rate in cancer patients, the identification of a set of clinical correlatives to CTCs in breast cancer is likely to facilitate the identification of those patients that would benefit the most from CTC analysis, including genetic profile assessment for patient stratification and testing of drug susceptibility. As an added benefit to this analysis, the identification of Denosumab treatment as a strategy to reduce CTC intravasation warrants further investigation.

Altogether, tremendous advancements in the CTC field have prepared the ground to investigate the functional features of cancer aggressiveness. Presented work profoundly extended the focus of CTC research and revealed new means for CTC characterization. Yet, the exact processes underlying cancer spread and metastasis vulnerability remain unknown and gaining more insights about the cancer perpetrators is still essential.

7. REFERENCES

1. Tomasetti, C. and B. Vogelstein, *Cancer etiology. Variation in cancer risk among tissues can be explained by the number of stem cell divisions.* Science, 2015. **347**(6217): p. 78-81.
2. Torre, L.A., et al., *Global cancer statistics, 2012.* CA Cancer J Clin, 2015. **65**(2): p. 87-108.
3. Jemal, A., et al., *Global cancer statistics.* CA Cancer J Clin, 2011. **61**(2): p. 69-90.
4. Talmadge, J.E. and I.J. Fidler, *AACR centennial series: the biology of cancer metastasis: historical perspective.* Cancer Res, 2010. **70**(14): p. 5649-69.
5. Massague, J. and A.C. Obenauf, *Metastatic colonization by circulating tumour cells.* Nature, 2016. **529**(7586): p. 298-306.
6. Yoo, B.H., et al., *Anoikis of colon carcinoma cells triggered by beta-catenin loss can be enhanced by tumor necrosis factor receptor 1 antagonists.* Oncogene, 2015. **34**(38): p. 4939-51.
7. Wheeler, L.J., et al., *CBX2 identified as driver of anoikis escape and dissemination in high grade serous ovarian cancer.* Oncogenesis, 2018. **7**(11): p. 92.
8. Vayrynen, S.A., et al., *Clinical impact and network of determinants of tumour necrosis in colorectal cancer.* Br J Cancer, 2016. **114**(12): p. 1334-42.
9. Bredholt, G., et al., *Tumor necrosis is an important hallmark of aggressive endometrial cancer and associates with hypoxia, angiogenesis and inflammation responses.* Oncotarget, 2015. **6**(37): p. 39676-91.
10. Kitamura, T., B.-Z. Qian, and J.W. Pollard, *Immune cell promotion of metastasis.* Nature reviews. Immunology, 2015. **15**(2): p. 73-86.
11. Chen, M.-T., et al., *Comparison of patterns and prognosis among distant metastatic breast cancer patients by age groups: a SEER population-based analysis.* Scientific reports, 2017. **7**(1): p. 9254-9254.
12. Valderrama-Treviño, A.I., et al., *Hepatic Metastasis from Colorectal Cancer.* Euroasian journal of hepato-gastroenterology, 2017. **7**(2): p. 166-175.
13. Ellis, P.M. and R. Vandermeer, *Delays in the diagnosis of lung cancer.* Journal of thoracic disease, 2011. **3**(3): p. 183-188.
14. Tivari, S., et al., *Reawakening of dormant estrogen-dependent human breast cancer cells by bone marrow stroma secretory senescence.* Cell Commun Signal, 2018. **16**(1): p. 48.
15. Lindstrom, L.S., et al., *Clinically used breast cancer markers such as estrogen receptor, progesterone receptor, and human epidermal growth factor receptor 2 are unstable throughout tumor progression.* J Clin Oncol, 2012. **30**(21): p. 2601-8.
16. Rybinski, B. and K. Yun, *Addressing intra-tumoral heterogeneity and therapy resistance.* Oncotarget, 2016. **7**(44): p. 72322-72342.
17. Rueda, O.M., et al., *Dynamics of breast-cancer relapse reveal late-recurring ER-positive genomic subgroups.* Nature, 2019.
18. Criscitiello, C., et al., *Biopsy confirmation of metastatic sites in breast cancer patients: clinical impact and future perspectives.* Breast cancer research : BCR, 2014. **16**(2): p. 205-205.
19. Toss, A., et al., *CTC enumeration and characterization: moving toward personalized medicine.* Ann Transl Med, 2014. **2**(11): p. 108.
20. Fidler, I.J., *The relationship of embolic homogeneity, number, size and viability to the incidence of experimental metastasis.* Eur J Cancer, 1973. **9**(3): p. 223-7.

21. Aceto, N., et al., *Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis.* Cell, 2014. **158**(5): p. 1110-1122.
22. Duda, D.G., et al., *Malignant cells facilitate lung metastasis by bringing their own soil.* Proc Natl Acad Sci U S A, 2010. **107**(50): p. 21677-82.
23. Sarioglu, A.F., et al., *A microfluidic device for label-free, physical capture of circulating tumor cell clusters.* Nat Methods, 2015. **12**(7): p. 685-91.
24. Yu, M., et al., *Circulating breast tumor cells exhibit dynamic changes in epithelial and mesenchymal composition.* Science, 2013. **339**(6119): p. 580-4.
25. Jansson, S., et al., *Prognostic impact of circulating tumor cell apoptosis and clusters in serial blood samples from patients with metastatic breast cancer in a prospective observational cohort.* BMC cancer, 2016. **16**: p. 433-433.
26. Aceto, N., et al., *En Route to Metastasis: Circulating Tumor Cell Clusters and Epithelial-to-Mesenchymal Transition.* Trends Cancer, 2015. **1**(1): p. 44-52.
27. Zheng, X., et al., *Epithelial-to-mesenchymal transition is dispensable for metastasis but induces chemoresistance in pancreatic cancer.* Nature, 2015. **527**(7579): p. 525-530.
28. Fischer, K.R., et al., *Epithelial-to-mesenchymal transition is not required for lung metastasis but contributes to chemoresistance.* Nature, 2015. **527**(7579): p. 472-6.
29. Gkountela, S., et al., *Recent advances in the biology of human circulating tumour cells and metastasis.* ESMO Open, 2016. **1**(4): p. e000078.
30. Maheswaran, S., et al., *Detection of mutations in EGFR in circulating lung-cancer cells.* N Engl J Med, 2008. **359**(4): p. 366-77.
31. Cristofanilli, M., et al., *The clinical use of circulating tumor cells (CTCs) enumeration for staging of metastatic breast cancer (MBC): International expert consensus paper.* Crit Rev Oncol Hematol, 2019. **134**: p. 39-45.
32. Joosse, S.A., T.M. Gorges, and K. Pantel, *Biology, detection, and clinical implications of circulating tumor cells.* EMBO Mol Med, 2015. **7**(1): p. 1-11.
33. Riethdorf, S., et al., *Detection of circulating tumor cells in peripheral blood of patients with metastatic breast cancer: a validation study of the CellSearch system.* Clin Cancer Res, 2007. **13**(3): p. 920-8.
34. Nagrath, S., et al., *Isolation of rare circulating tumour cells in cancer patients by microchip technology.* Nature, 2007. **450**(7173): p. 1235-9.
35. Ozkumur, E., et al., *Inertial focusing for tumor antigen-dependent and -independent sorting of rare circulating tumor cells.* Sci Transl Med, 2013. **5**(179): p. 179ra47.
36. Chudziak, J., et al., *Clinical evaluation of a novel microfluidic device for epitope-independent enrichment of circulating tumour cells in patients with small cell lung cancer.* Analyst, 2016. **141**(2): p. 669-78.
37. Hou, H.W., et al., *Isolation and retrieval of circulating tumor cells using centrifugal forces.* Sci Rep, 2013. **3**: p. 1259.
38. Galanzha, E.I. and V.P. Zharov, *Circulating Tumor Cell Detection and Capture by Photoacoustic Flow Cytometry in Vivo and ex Vivo.* Cancers (Basel), 2013. **5**(4): p. 1691-738.
39. Tauriello, D.V.F., et al., *TGFbeta drives immune evasion in genetically reconstituted colon cancer metastasis.* Nature, 2018. **554**(7693): p. 538-543.
40. Quail, D.F. and J.A. Joyce, *Microenvironmental regulation of tumor progression and metastasis.* Nature medicine, 2013. **19**(11): p. 1423-1437.
41. Salvatore, V., et al., *The tumor microenvironment promotes cancer progression and cell migration.* Oncotarget, 2016. **8**(6): p. 9608-9616.

42. Mlecnik, B., et al., *The tumor microenvironment and Immunoscore are critical determinants of dissemination to distant metastasis.* Sci Transl Med, 2016. **8**(327): p. 327ra26.
43. Wculek, S.K. and I. Malanchi, *Neutrophils support lung colonization of metastasis-initiating breast cancer cells.* Nature, 2015. **528**(7582): p. 413-7.
44. Toor, S.M., et al., *Myeloid cells in circulation and tumor microenvironment of breast cancer patients.* Cancer Immunol Immunother, 2017. **66**(6): p. 753-764.
45. Li, B., et al., *Landscape of tumor-infiltrating T cell repertoire of human cancers.* Nat Genet, 2016. **48**(7): p. 725-32.
46. Soncin, I., et al., *The tumour microenvironment creates a niche for the self-renewal of tumour-promoting macrophages in colon adenoma.* Nat Commun, 2018. **9**(1): p. 582.
47. Michea, P., et al., *Adjustment of dendritic cells to the breast-cancer microenvironment is subset specific.* Nat Immunol, 2018. **19**(8): p. 885-897.
48. Sjodahl, G., et al., *Infiltration of CD3(+) and CD68(+) cells in bladder cancer is subtype specific and affects the outcome of patients with muscle-invasive tumors.* Urol Oncol, 2014. **32**(6): p. 791-7.
49. Sideras, K., et al., *PD-L1, Galectin-9 and CD8(+) tumor-infiltrating lymphocytes are associated with survival in hepatocellular carcinoma.* Oncoimmunology, 2017. **6**(2): p. e1273309.
50. Hotta, K., et al., *Magnitude of the benefit of progression-free survival as a potential surrogate marker in phase 3 trials assessing targeted agents in molecularly selected patients with advanced non-small cell lung cancer: systematic review.* PLoS One, 2015. **10**(3): p. e0121211.
51. He, G., et al., *Peritumoural neutrophils negatively regulate adaptive immunity via the PD-L1/PD-1 signalling pathway in hepatocellular carcinoma.* J Exp Clin Cancer Res, 2015. **34**: p. 141.
52. Feng, S., et al., *Myeloid-derived suppressor cells inhibit T cell activation through nitrating LCK in mouse cancers.* Proc Natl Acad Sci U S A, 2018. **115**(40): p. 10094-10099.
53. Su, S., et al., *Immune Checkpoint Inhibition Overcomes ADCP-Induced Immunosuppression by Macrophages.* Cell, 2018. **175**(2): p. 442-457.e23.
54. Multhoff, G., M. Molls, and J. Radons, *Chronic inflammation in cancer development.* Front Immunol, 2011. **2**: p. 98.
55. Stockmann, C., et al., *Deletion of vascular endothelial growth factor in myeloid cells accelerates tumorigenesis.* Nature, 2008. **456**(7223): p. 814-8.
56. Wilke, C.M., et al., *Prognostic significance of regulatory T cells in tumor.* Int J Cancer, 2010. **127**(4): p. 748-58.
57. Wolf, A.M., et al., *Increase of regulatory T cells in the peripheral blood of cancer patients.* Clin Cancer Res, 2003. **9**(2): p. 606-12.
58. Kerkar, S.P. and N.P. Restifo, *Cellular constituents of immune escape within the tumor microenvironment.* Cancer Res, 2012. **72**(13): p. 3125-30.
59. Gong, J., et al., *Development of PD-1 and PD-L1 inhibitors as a form of cancer immunotherapy: a comprehensive review of registration trials and future considerations.* J Immunother Cancer, 2018. **6**(1): p. 8.
60. Rowshanravan, B., N. Halliday, and D.M. Sansom, *CTLA-4: a moving target in immunotherapy.* Blood, 2018. **131**(1): p. 58-67.
61. Du, W., et al., *TIM-3 as a Target for Cancer Immunotherapy and Mechanisms of Action.* Int J Mol Sci, 2017. **18**(3).

62. Powell, A.A., et al., *Single cell profiling of circulating tumor cells: transcriptional heterogeneity and diversity from breast cancer cell lines.* PLoS One, 2012. **7**(5): p. e33788.
63. Polzer, B., et al., *Molecular profiling of single circulating tumor cells with diagnostic intention.* EMBO Mol Med, 2014. **6**(11): p. 1371-86.
64. Suzuki, A., et al., *Single-cell analysis of lung adenocarcinoma cell lines reveals diverse expression patterns of individual cells invoked by a molecular target drug treatment.* Genome Biol, 2015. **16**: p. 66.
65. Macaulay, I.C., et al., *G&T-seq: parallel sequencing of single-cell genomes and transcriptomes.* Nat Methods, 2015. **12**(6): p. 519-22.
66. Schwartzman, O. and A. Tanay, *Single-cell epigenomics: techniques and emerging applications.* Nat Rev Genet, 2015. **16**(12): p. 716-26.
67. Wu, M. and A.K. Singh, *Single-cell protein analysis.* Curr Opin Biotechnol, 2012. **23**(1): p. 83-8.
68. Yu, M., et al., *Cancer therapy. Ex vivo culture of circulating breast tumor cells for individualized testing of drug susceptibility.* Science, 2014. **345**(6193): p. 216-20.
69. Macaulay, I.C., et al., *Separation and parallel sequencing of the genomes and transcriptomes of single cells using G&T-seq.* Nat Protoc, 2016. **11**(11): p. 2081-103.
70. Chen, S., et al., *Genome-wide CRISPR screen in a mouse model of tumor growth and metastasis.* Cell, 2015. **160**(6): p. 1246-60.
71. Li, H., et al., *Reference component analysis of single-cell transcriptomes elucidates cellular heterogeneity in human colorectal tumors.* Nat Genet, 2017. **49**(5): p. 708-718.
72. Szczerba, B.M., et al., *Neutrophils escort circulating tumour cells to enable cell cycle progression.* Nature, 2019. **566**(7745): p. 553-557.
73. Pantel, K. and M.R. Speicher, *The biology of circulating tumor cells.* Oncogene, 2016. **35**(10): p. 1216-24.
74. Xu, L., et al., *Optimization and Evaluation of a Novel Size Based Circulating Tumor Cell Isolation System.* PLoS One, 2015. **10**(9): p. e0138032.
75. Fridlender, Z.G., et al., *Polarization of tumor-associated neutrophil phenotype by TGF-beta: "N1" versus "N2" TAN.* Cancer Cell, 2009. **16**(3): p. 183-94.
76. Cristofanilli, M., et al., *Circulating tumor cells, disease progression, and survival in metastatic breast cancer.* N Engl J Med, 2004. **351**(8): p. 781-91.
77. Labelle, M., S. Begum, and R.O. Hynes, *Direct signaling between platelets and cancer cells induces an epithelial-mesenchymal-like transition and promotes metastasis.* Cancer Cell, 2011. **20**(5): p. 576-90.
78. Rothstein, G., et al., *Stimulation of neutrophil production in CSF-1-responsive clones.* Blood, 1988. **72**(3): p. 898-902.
79. He, J.Q., et al., *Association of genetic variations in the CSF2 and CSF3 genes with lung function in smoking-induced COPD.* Eur Respir J, 2008. **32**(1): p. 25-34.
80. Verri, W.A., Jr., et al., *IL-15 mediates antigen-induced neutrophil migration by triggering IL-18 production.* Eur J Immunol, 2007. **37**(12): p. 3373-80.
81. Kacinski, B.M., et al., *FMS (CSF-1 receptor) and CSF-1 transcripts and protein are expressed by human breast carcinomas in vivo and in vitro.* Oncogene, 1991. **6**(6): p. 941-52.
82. McCracken, J.M. and L.A. Allen, *Regulation of human neutrophil apoptosis and lifespan in health and disease.* J Cell Death, 2014. **7**: p. 15-23.
83. Canli, O., et al., *Myeloid Cell-Derived Reactive Oxygen Species Induce Epithelial Mutagenesis.* Cancer Cell, 2017. **32**(6): p. 869-883 e5.

84. Alexandrov, L.B., et al., *Signatures of mutational processes in human cancer.* Nature, 2013. **500**(7463): p. 415-21.
85. Wellenstein, M.D. and K.E. de Visser, *Cancer-Cell-Intrinsic Mechanisms Shaping the Tumor Immune Landscape.* Immunity, 2018. **48**(3): p. 399-416.
86. Ramasamy, S., et al., *Tle1 tumor suppressor negatively regulates inflammation in vivo and modulates NF-kappaB inflammatory pathway.* Proc Natl Acad Sci U S A, 2016. **113**(7): p. 1871-6.
87. Egeblad, M. and K.E. de Visser, *Sticking together helps cancer to spread.* Nature, 2019. **566**(7745): p. 459-460.
88. Vetter, M., et al., *Denosumab treatment is associated with the absence of circulating tumor cells in patients with breast cancer.* Breast Cancer Res, 2018. **20**(1): p. 141.
89. Rack, B., et al., *Circulating tumor cells predict survival in early average-to-high risk breast cancer patients.* J Natl Cancer Inst, 2014. **106**(5).
90. Giuliano, M., et al., *Circulating tumor cells as prognostic and predictive markers in metastatic breast cancer patients receiving first-line systemic treatment.* Breast Cancer Res, 2011. **13**(3): p. R67.
91. Carter, L., et al., *Molecular analysis of circulating tumor cells identifies distinct copy-number profiles in patients with chemosensitive and chemorefractory small-cell lung cancer.* Nat Med, 2017. **23**(1): p. 114-119.
92. Kaifi, J.T., et al., *Circulating tumor cell levels are elevated in colorectal cancer patients with high tumor burden in the liver.* Cancer Biol Ther, 2015. **16**(5): p. 690-8.
93. Cristofanilli, M., et al., *Circulating tumor cells in metastatic breast cancer: biologic staging beyond tumor burden.* Clin Breast Cancer, 2007. **7**(6): p. 471-9.
94. Whitaker, M., et al., *Bisphosphonates for osteoporosis--where do we go from here?* N Engl J Med, 2012. **366**(22): p. 2048-51.
95. Hanley, D.A., et al., *Denosumab: mechanism of action and clinical outcomes.* Int J Clin Pract, 2012. **66**(12): p. 1139-46.
96. Blake, M.L., et al., *RANK expression on breast cancer cells promotes skeletal metastasis.* Clin Exp Metastasis, 2014. **31**(2): p. 233-45.
97. Pfitzner, B.M., et al., *RANK expression as a prognostic and predictive marker in breast cancer.* Breast Cancer Res Treat, 2014. **145**(2): p. 307-15.
98. Jordan, N.V., et al., *HER2 expression identifies dynamic functional states within circulating breast cancer cells.* Nature, 2016. **537**(7618): p. 102-106.
99. Houssami, N., et al., *HER2 discordance between primary breast cancer and its paired metastasis: tumor biology or test artefact? Insights through meta-analysis.* Breast Cancer Res Treat, 2011. **129**(3): p. 659-74.
100. Anderson, W.F., et al., *Comparison of age distribution patterns for different histopathologic types of breast carcinoma.* Cancer Epidemiol Biomarkers Prev, 2006. **15**(10): p. 1899-905.
101. Adami, H.O., et al., *The relation between survival and age at diagnosis in breast cancer.* N Engl J Med, 1986. **315**(9): p. 559-63.
102. Fredholm, H., et al., *Long-term outcome in young women with breast cancer: a population-based study.* Breast Cancer Res Treat, 2016. **160**(1): p. 131-143.
103. Bleyer, A., et al., *The distinctive biology of cancer in adolescents and young adults.* Nat Rev Cancer, 2008. **8**(4): p. 288-98.
104. Koleckova, M., et al., *Age-associated prognostic and predictive biomarkers in patients with breast cancer.* Oncol Lett, 2017. **13**(6): p. 4201-4207.
105. Duffy, M.J., *CA 15-3 and related mucins as circulating markers in breast cancer.* Ann Clin Biochem, 1999. **36 (Pt 5)**: p. 579-86.

106. Oken, M.M., et al., *Toxicity and response criteria of the Eastern Cooperative Oncology Group.* Am J Clin Oncol, 1982. **5**(6): p. 649-55.
107. Khoo, B.L., et al., *Expansion of patient-derived circulating tumor cells from liquid biopsies using a CTC microfluidic culture device.* Nat Protoc, 2018. **13**(1): p. 34-58.
108. Lu, X. and Y. Kang, *Hypoxia and hypoxia-inducible factors: master regulators of metastasis.* Clin Cancer Res, 2010. **16**(24): p. 5928-35.
109. Verma, R., et al., *Lymphocyte depletion and repopulation after chemotherapy for primary breast cancer.* Breast cancer research : BCR, 2016. **18**(1): p. 10-10.
110. Grimm, R.H., Jr., J.D. Neaton, and W. Ludwig, *Prognostic importance of the white blood cell count for coronary, cancer, and all-cause mortality.* Jama, 1985. **254**(14): p. 1932-7.
111. Bruckner, H.W., et al., *Absolute Granulocyte, Lymphocyte, and Monocyte Counts: Useful Determinants of Prognosis for Patients With Metastatic Cancer of the Stomach.* JAMA, 1982. **247**(7): p. 1004-1006.
112. Aliustaoglu, M., et al., *The effect of peripheral blood values on prognosis of patients with locally advanced gastric cancer before treatment.* 2010. **27**(4): p. 1060-1065.
113. Sasaki, A., et al., *Prognostic Value of Preoperative Peripheral Blood Monocyte Count in Patients with Colorectal Liver Metastasis after Liver Resection.* 2007. **11**(5): p. 596-602.
114. Gkountela, S., et al., *Circulating Tumor Cell Clustering Shapes DNA Methylation to Enable Metastasis Seeding.* Cell, 2019. **176**(1-2): p. 98-112 e14.
115. Krol, I., et al., *Detection of circulating tumour cell clusters in human glioblastoma.* Br J Cancer, 2018. **119**(4): p. 487-491.

OTHER SCIENTIFIC CONTRIBUTIONS

Over the course of my studies I had a chance to support multiple other projects, all listed below.

Internal:

Dr. Sofia Gkountela on "Circulating Tumor Cell Clustering Shapes DNA Methylation to Enable Metastasis Seeding", Cell, 2019 [114]

Ilona Krol on "Detection of circulating tumour cell clusters in human glioblastoma", British Journal of Cancer, 2018 [115]

Cinzia Donato on "Hypoxia Triggers the Intravasation of Clustered Circulating Tumor Cells", ongoing PhD project

Manuel Scheidmann on "Genome-wide CRISPR Screen to Identify Genes Required for CTC Cluster Formation and Maintenance", ongoing PhD project

Aino Alise Paasinen Sohns on "Tracing the regions of origins of various liquid biopsy components circulating in blood using barcoded human breast cancer xenograft models", ongoing project

External:

Prof. Dr. Andrew Ewald on "E-cadherin is an invasion suppressor, survival factor, and metastasis promoter across multiple models of breast cancer", Nature, under revisions

Dr. Marta Cavo on "A novel bioreactor enables spontaneous metastasis modeling in vitro", Biotechnology and Bioengineering, manuscript in preparation

Stefanie Tiede on "Tumor heterogeneity during the progression of metastatic breast cancer", a PhD project

Fabiana Lüönd on "Tracing Epithelial-to-Mesenchymal Transition in Breast Cancer", a PhD project

Dr. Nami Matsuda, Dr. Vida Vafaizadeh and Prof. Dr. Maries van den Broek on CTC detection from mouse models

Dr. Alessio Zippo on sequencing of CTC and primary tumor in NSG-IMEC-MYC-PIK3CA model

Prof. Dr. Luigi Mariani on establishing single cell parallel DNA and RNA sequencing

Armbrecht Lucas on establishing microfluidic chip for CTC capture and characterization

Team of Prof. Dr. Timm Schroeder on hematopoietic cells isolation including semi-adherent cells

Dr. Roberta Carbone from Tethis on KRAS sequencing from laser microdissected CTCs

Prof. Dr. Werner Baumgartner on *in vivo* CTC cluster dissociation

PD Dr. Andrea Banfi on picking fixed tissue from cryosections and RNA-seq

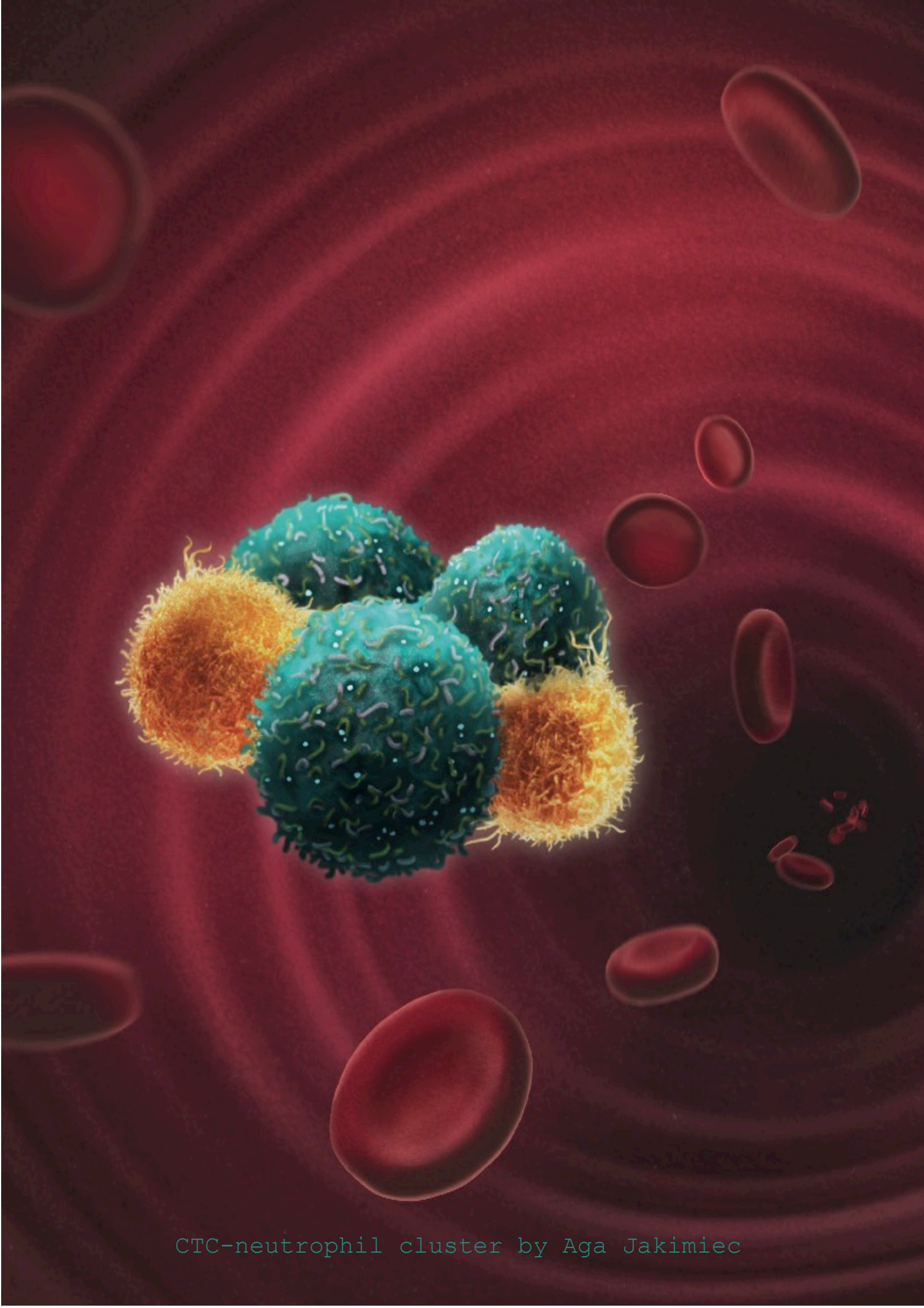
Prof. Tobias Junt from Novartis on his Sjogren's Syndrome studies

Moreover, I have contributed by co-authoring several review articles:

Gkountela S, **Szczerba B**, Donato C, Aceto N. Recent advances in the biology of human circulating tumour cells and metastasis. ESMO Open, 2016 [29]

Marass F, Castro-Giner F, **Szczerba B**, Jahn K, Kuipers J, Aceto N, Beerenwinkel N. Computational analysis of DNA and RNA sequencing data obtained from liquid biopsies. Book chapter in 'Liquid Tumor Biopsies' by Springer, 2019 (in press)

CO-AUTHORED PUBLICATIONS



CTC-neutrophil cluster by Aga Jakimiec

LETTER

https://doi.org/10.1038/s41586-019-0915-y

Neutrophils escort circulating tumour cells to enable cell cycle progression

Barbara Maria Szczerba[1], Francesc Castro-Giner[1,2], Marcus Vetter[3,4], Ilona Krol[1], Sofia Gkountela[1], Julia Landin[4], Manuel C. Scheidmann[1], Cinzia Donato[1], Ramona Scherrer[1], Jochen Singer[2,5], Christian Beisel[5], Christian Kurzeder[3,6], Viola Heinzelmann-Schwarz[3], Christoph Rochlitz[4], Walter Paul Weber[6], Niko Beerenwinkel[2,5] & Nicola Aceto[1]*

A better understanding of the features that define the interaction between cancer cells and immune cells is important for the development of new cancer therapies[1]. However, focus is often given to interactions that occur within the primary tumour and its microenvironment, whereas the role of immune cells during cancer dissemination in patients remains largely uncharacterized[2,3]. Circulating tumour cells (CTCs) are precursors of metastasis in several types of cancer[4–6], and are occasionally found within the bloodstream in association with non-malignant cells such as white blood cells (WBCs)[7,8]. The identity and function of these CTC-associated WBCs, as well as the molecular features that define the interaction between WBCs and CTCs, are unknown. Here we isolate and characterize individual CTC-associated WBCs, as well as corresponding cancer cells within each CTC–WBC cluster, from patients with breast cancer and from mouse models. We use single-cell RNA sequencing to show that in the majority of these cases, CTCs were associated with neutrophils. When comparing the transcriptome profiles of CTCs associated with neutrophils against those of CTCs alone, we detect a number of differentially expressed genes that outline cell cycle progression, leading to more efficient metastasis formation. Further, we identify cell–cell junction and cytokine–receptor pairs that define CTC–neutrophil clusters, representing key vulnerabilities of the metastatic process. Thus, the association between neutrophils and CTCs drives cell cycle progression within the bloodstream and expands the metastatic potential of CTCs, providing a rationale for targeting this interaction in treatment of breast cancer.

CTCs are precursors of metastasis in various solid cancers, including breast cancer[6], and are occasionally found in association with WBCs[7]. The role of CTC–WBC clusters in the development of metastasis as well as the principles that govern the interactions between CTCs and WBCs during blood-borne dissemination are largely uncharacterized.

We first sought to determine the number and composition of CTC–WBC clusters in patients with breast cancer and mouse models. We obtained blood samples from 70 patients with invasive breast cancer who discontinued their treatment owing to progressive disease, as well as from five different breast cancer mouse models, and enriched for CTCs using the Parsortix microfluidic device[9] (Extended Data Fig. 1a–e). We stained CTCs for cancer-associated cell-surface markers EpCAM, HER2 and EGFR or imaged directly for the expression of GFP, and labelled for CD45 to identify WBCs (Fig. 1a, Extended Data Fig. 1f). Among 70 patients, 34 (48.6%) had detectable CTCs, with a mean number of 22 CTCs per 7.5 ml of blood (Supplementary Tables 1, 2). Whereas the majority of CTCs were single (88.0%), we also detected CTC clusters (8.6%) and CTC–WBC clusters (3.4%) (Fig. 1b, Extended Data Fig. 1g, h). Similarly, we observed that CTC–WBC clusters were present in all of the mouse models we tested, including those with immunodeficient or immunocompetent backgrounds, ranging from 0.05% to 61% of the total CTC population (Fig. 1b, Extended Data Fig. 1i, j). Of note, CTC abundance and ratios markedly changed when drawing blood upstream of capillary beds as opposed to locations further downstream, which indicates that clustered CTCs are shed early yet may be trapped in capillaries before reaching the periphery (Extended Data Fig. 1k–n). Thus, CTC–WBC clusters are rare in the peripheral circulation yet consistently found in patients with breast cancer and in mouse models.

We then determined the type of WBCs found in CTC–WBC clusters. We used a robotic micromanipulator to dissociate CTC–WBC clusters (Supplementary Video 1), enabling single-cell RNA sequencing analysis of cluster-associated WBCs and their comparison to reference WBCs from matched donors (Fig. 1c) using reference component analysis[10]. In patients, we found that 75% of CTC-associated WBCs were related to the myeloid lineage, whereas the remaining (25%) were similar to T cells (Fig. 1d, Extended Data Fig. 2a, b). Likewise, we found that 93% of CTC-associated WBCs from mouse models were also characterized by a myeloid-cell-like expression profile (Extended Data Fig. 2c–e). To determine the proportion of CTC-associated WBCs that are neutrophils, monocytes or macrophages, we labelled CTC–WBC clusters for Ly-6G, CD11b and F4/80, as well as with Wright–Giemsa staining to define nuclear morphology (Extended Data Fig. 2f, g). We found that the vast majority (85.5%–91.7%) of CTC-associated WBCs were positive for Ly-6G and displayed a nuclear morphology typical of neutrophils, whereas a minority (8.3%–14.5%) were monocytes ($CD11b^+F4/80^-Ly\text{-}6G^-$), and no $F4/80^+$ macrophages were found (Fig. 1e, f, Extended Data Fig. 2h–j). Further, RNA sequencing analysis revealed expression of *ARG1*, *CXCL1*, *CXCL2*, *CXCL10*, *CCL2*, *CXCR2* and *VEGFA* in most CTC-associated neutrophils from patients (or their orthologues in mouse models) (Extended Data Fig. 2k), which indicates that CTC-associated neutrophils have a gene expression profile that is similar to that of pro-tumour N2-like cells[11].

We next investigated whether the presence of CTC–neutrophil clusters in patients with breast cancer could predict disease outcome. Of note, patients in whom at least one CTC–neutrophil cluster was detected in 7.5 ml of peripheral blood were characterized by a significantly worse progression-free survival compared to patients with five or more CTCs per 7.5 ml of peripheral blood (previously defined as a threshold for adverse outcome[12]) (Fig. 1g), as well as when compared to all patients with no CTC–neutrophil clusters, patients with at least one CTC per 7.5 ml of blood or patients in whom either single CTCs or CTC clusters were found (Extended Data Fig. 3a–c). In addition, we individually micromanipulated equal numbers of CTCs from CTC–neutrophil clusters, CTC clusters and single CTCs, spontaneously generated from tumour-bearing mice, and intravenously injected 100 CTCs per mouse into tumour-free recipient mice from each of these categories. We found that mice injected with CTCs from CTC–neutrophil clusters developed overt metastasis much faster than those injected with CTCs alone and, accordingly, survived for a shorter time (Fig. 1h, Extended Data Fig. 3d–h). Thus, CTC–neutrophil clusters represent

[1]Department of Biomedicine, Cancer Metastasis Lab, University of Basel and University Hospital Basel, Basel, Switzerland. [2]SIB Swiss Institute of Bioinformatics, Lausanne, Switzerland. [3]Gynecologic Cancer Center, University Hospital Basel, Basel, Switzerland. [4]Department of Medical Oncology, University Hospital Basel, Basel, Switzerland. [5]Department of Biosystems Science and Engineering, ETH Zurich, Basel, Switzerland. [6]Breast Center, University of Basel and University Hospital Basel, Basel, Switzerland. *e-mail: Nicola.Aceto@unibas.ch

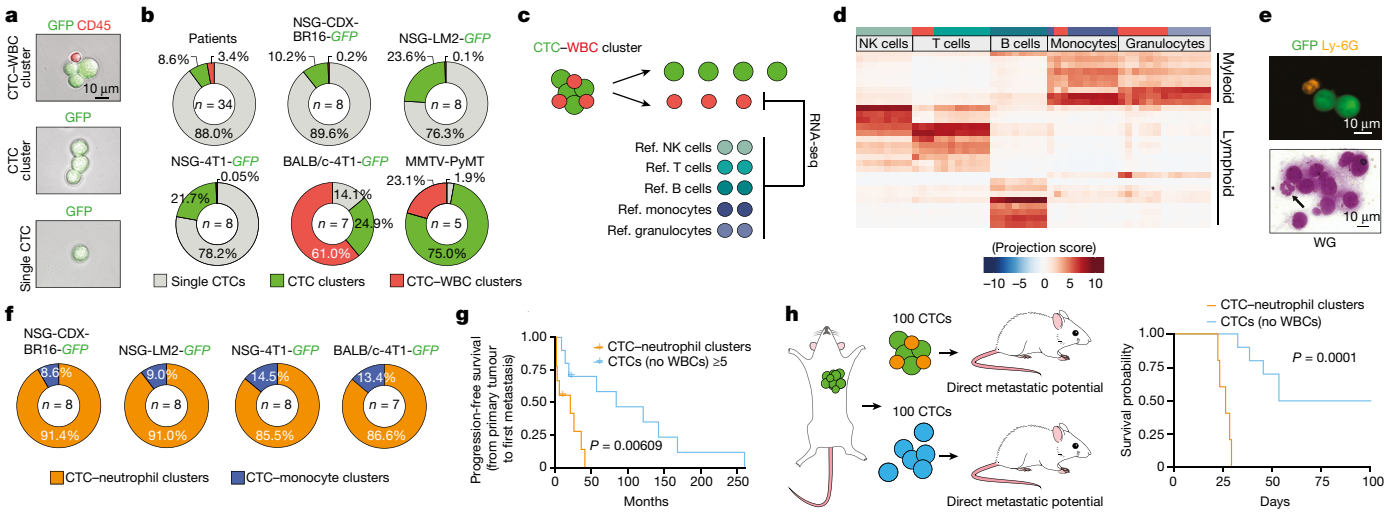


Fig. 1 | CTC–neutrophil clusters are highly efficient metastatic precursors. a, Representative images of a CTC–WBC cluster, a CTC cluster and a single CTC from NSG-CDX-BR16-*GFP* mice. CTCs are expressing GFP (green) and CTC-associated WBCs are labelled with anti-CD45 antibodies (red) ($n = 8$). **b**, Pie charts displaying the mean percentage of single CTCs (grey), CTC clusters (green) and CTC–WBC clusters (red) in patients with breast cancer and in mouse models. The number of independent biological replicates (n) is shown for each model. **c**, Schematic of the experimental design. CTC–WBC clusters are dissociated into individual cells and processed for RNA sequencing (RNA-seq). The transcriptome of CTC-associated WBCs is then compared to reference (ref.) WBCs. **d**, Reference component analysis clustering of CTC-associated WBCs and reference WBCs from patients with breast cancer, displaying projection scores of cells (columns; $n = 50$) on the immune reference panel (rows). **e**, Representative images of CTC–WBC clusters stained for Ly-6G (neutrophils, gold) together with GFP (cancer cells, green) (top) or processed with the Wright–Giemsa (WG) assay to define nuclear morphology (bottom) ($n = 8$). **f**, Pie charts showing the mean percentage of CTC–neutrophil clusters and CTC–monocyte clusters in individual models. The number of independent biological replicates (n) is shown for each model. **g**, Kaplan–Meier plot showing progression-free survival of patients with breast cancer ($n = 9$ for patients with one or more CTC–neutrophil clusters and $n = 10$ for patients with five or more CTCs); P value by two-sided log-rank test is shown. **h**, Schematic of the experimental design. One hundred CTCs from CTC–neutrophil clusters or CTCs alone are injected in the tail vein of recipient mice to measure their metastatic potential (left). Kaplan–Meier plot showing overall survival of mice (right). $n = 5$ for CTC–neutrophil clusters and $n = 10$ for CTCs alone; P value by two-sided log-rank test is shown.

the most efficient metastasis-forming cell subpopulation among breast CTCs, and their presence in the bloodstream of patients is associated with a poor prognosis.

We next sought to determine the molecular outcome of the interaction between CTCs and neutrophils by dissociating CTC–neutrophil clusters and comparing the expression profile of CTCs from CTC–neutrophil clusters to that of CTCs alone (Fig. 2a). We first determined differential gene expression in the syngeneic BALB/c-4T1-*GFP* model, which provided the highest number of CTCs from CTC–neutrophil clusters ($n = 25$). Compared to CTCs that were not associated with neutrophils, we found that CTCs from CTC–neutrophil clusters are characterized by differential expression of a set of 51 genes, of which 41 are upregulated and 10 are downregulated (Fig. 2b, Supplementary Tables 3, 4). Pathway analysis of the upregulated genes revealed that CTCs from CTC–neutrophil clusters display a marked enrichment in positive regulators of cell cycle and DNA replication programs compared to CTCs alone (Fig. 2c), independently of the number of detected features or reads in each individual sample (Extended Data Fig. 4a). The same results were obtained with CTCs from patients with breast cancer (Fig. 2c, Extended Data Fig. 4b). Immunofluorescence staining of CTCs confirmed higher levels of Ki67 expression in CTCs from CTC–neutrophil clusters (Fig. 2d, e), consistent with the RNA sequencing results. By contrast, no significant differences were observed for genes involved in epithelial-to-mesenchymal transition, cancer stem-cell markers or platelet-related genes[13] (Extended Data Fig. 4c–h). We further investigated whether neutrophil proximity confers a proliferative advantage to cancer cells at the level of the primary tumour, disseminated tumour cells and overt metastasis. We found that Ki67 expression does not increase in cancer cells that surround tumour-infiltrated neutrophils in the primary tumour or overt metastasis (Extended Data Fig. 5a–d). However, higher Ki67 expression is retained in disseminated tumour cells from CTC–neutrophil clusters (Extended Data Fig. 5e, f), that is, when they are deprived of other stromal-derived signals that are typical of overt disease.

We then investigated which cytokines are expressed by CTC-associated neutrophils and paralleled by simultaneous expression of matching cytokine receptor(s) in the corresponding cancer cells. We found that genes that encode four cytokines (TNF-α, OSM, IL-1β and IL-6) are most frequently expressed by CTC-associated neutrophils of patients or patient-derived mouse models and matched by expression of their receptors on the corresponding cancer cells (Extended Data Fig. 6a). With a reverse approach, we also found that CTCs from CTC–neutrophil clusters most frequently expressed genes that encode cytokines such as CSF1, CSF3 (also known as granulocyte colony-stimulating factor (G-CSF)), TGF-β3 and IL-15, possibly involved in neutrophil stimulation[14–17], whereas corresponding neutrophils expressed their receptors (Extended Data Fig. 6b). Consistently, we observed that a 24-h in vitro treatment (coherent with neutrophil lifespan[18]) with IL6, IL-1β or both was sufficient to confer proliferative advantage to 4T1 cells upon dissemination, leading to faster metastasis development and shorter overall survival of mice (Fig. 2f, g, Extended Data Fig. 6c–e). Further, CRISPR-mediated knockout of IL-6 or IL-1β receptors in cancer cells (namely, *IL6ST* and *IL1R1*) did not alter the frequency of spontaneously generated CTC–neutrophil clusters, but did suppress their proliferative advantage (Extended Data Fig. 6f–h).

Given recent findings that highlight that the presence of myeloid cells in the primary tumour site leads to accumulation of mutational events[19], we investigated whether the mutational load of CTCs obtained from CTC–neutrophil clusters is different from that of CTCs alone in patients (Fig. 3a). Interestingly, we found that the mutational burden is similar between CTCs isolated from CTC–neutrophil clusters and CTCs alone, as well as when comparing donors with or without CTC–neutrophil clusters (Extended Data Fig. 7a, b). However, we observed that the overall frequency of C>T mutations was increased in CTCs from CTC–neutrophil clusters compared to CTCs alone, as

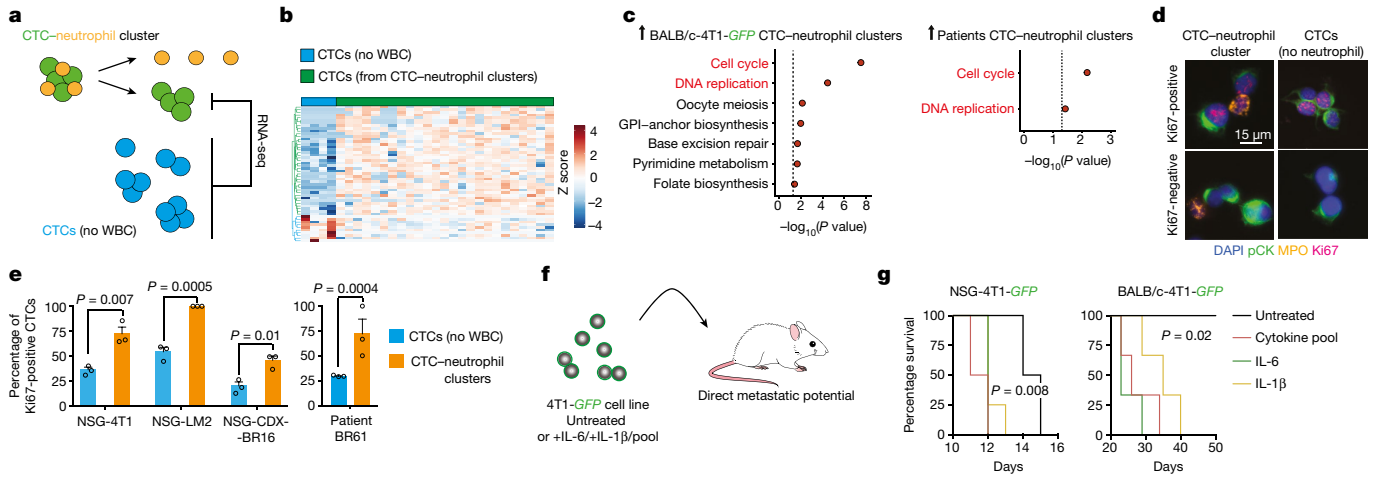


Fig. 2 | CTCs from CTC–neutrophil clusters are highly proliferative. **a**, Schematic of the experimental design. CTC–neutrophil clusters are dissociated and then processed for RNA sequencing. The transcriptome of CTCs from CTC–neutrophil clusters is compared to that of CTCs alone. **b**, Heat map of genes differentially expressed between CTCs from CTC–neutrophil clusters ($n = 25$) and CTCs alone ($n = 4$), isolated from BALB/c-4T1-*GFP* mice. The heat map displays gene-scaled (z-score) $\log_2$ counts per million mapped reads values after normalization, with columns representing samples ($n = 29$) and rows representing genes ($q < 0.05$ by edgeR likelihood ratio test). **c**, KEGG pathways over-represented ($P < 0.05$ by one-sided hypergeometric test) among upregulated genes in CTCs of CTC–neutrophil clusters from BALB/c-4T1-*GFP* mice (left) and test of selected pathways in patients ($P < 0.05$ by rotation gene set test (ROAST); right). **d**, Representative pictures of CTC–neutrophil clusters and CTCs alone from NSG-CDX-BR16-*GFP* mice, stained for pan-cytokeratin (pCK, green), myeloperoxidase (MPO, gold), Ki67 (purple) and DAPI (nuclei, blue) ($n = 3$). **e**, Plots showing the mean percentage of Ki67-positive CTCs. $n = 3$ for all; error bars, s.e.m.; P values by two-sided Student's t-test are shown. **f**, Schematic of the experimental design. 4T1-*GFP* cells are stimulated for 24 h with IL6, IL-1β or both (pool), then injected in recipient mice to assess their metastatic potential. **g**, Kaplan–Meier survival analysis of NSG (left) or BALB/c (right) mice injected with cytokine-treated 4T1-*GFP* cells. $n = 4$ for NSG-4T1-*GFP*, $n = 3$ for BALB/c-4T1-*GFP*, P values by two-sided log-rank test are shown.

well as in donors with CTC–neutrophil clusters, independent of the nucleotide context (that is, the nucleotides surrounding the mutation site) (Fig. 3b, Extended Data Fig. 7c, d). Although a general age-related accumulation of C>T mutations has previously been reported[20], we did not observe any age difference between the two groups (Extended Data Fig. 7e). Next, considering only high-impact mutations as defined by the genomic variant annotations and functional effect prediction toolbox (SnpEff), we asked whether specific genes are exclusively and recurrently mutated in donors with CTC–neutrophil clusters. This scenario is consistent with a model in which particular genetic alterations influence the recruitment of immune cells to the primary tumour[21] and increase the likelihood of generating CTC–neutrophil clusters. We found that a number of genes are indeed carrying high-impact mutations exclusively in donors with CTC–neutrophil clusters (Extended Data Fig. 7f, Supplementary Table 5). We then engineered 4T1-*GFP* cells to individually express all of the mutations found in two of the most frequently mutated genes (*MERTK* and *TLE1*) and injected them in the mammary gland of NSG (NOD-*scid*-*Il2rg*^{null}) mice. We observed that the introduction of *TLE1* mutations G1787A or G1509C led to a higher infiltration of neutrophils in the primary tumour and a higher proportion of spontaneously generated CTC–neutrophil clusters (33–41-fold increase), without affecting primary tumour size (Fig. 3c, Extended Data Fig. 7g–k). These results are consistent with recent observations involving *TLE1* function in regulating infiltration of myeloid cells into normal and neoplastic tissues[22]. Conversely, co-culture of cancer cells with neutrophils did not result in the acquisition of mutations within the same hotspots (Extended Data Fig. 8a, b). Together, our data reveals that although the overall mutational load remains unchanged, donors with CTC–neutrophil clusters feature a higher proportion of C>T substitutions and the presence of high-impact recurrent mutations in genes that promote neutrophil recruitment, such as *TLE1*.

We next tested whether the depletion or augmentation of the total neutrophil population would affect the ratio of spontaneously generated CTC–neutrophil clusters. We depleted neutrophils through in vivo

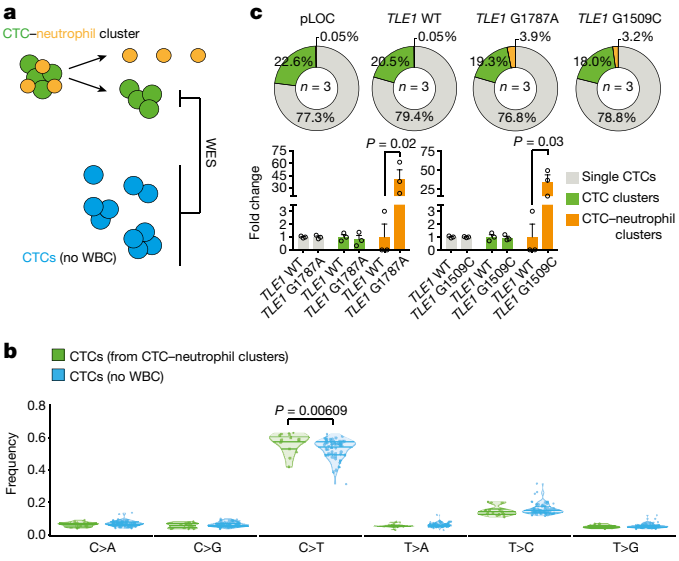


Fig. 3 | Whole-exome sequencing highlights recurrent mutational events in CTCs from CTC–neutrophil clusters. **a**, Schematic of the experimental design. CTC–neutrophil clusters are dissociated and then processed for whole-exome sequencing (WES). CTCs that were associated to neutrophils are compared to CTCs alone. **b**, Mutation distribution in CTCs from CTC–neutrophil clusters ($n = 14$) compared to CTCs alone ($n = 56$). Lines within the violin plots show the 25th, 50th and 75th percentile, respectively, and dots represent individual CTCs. P value by two-sided Wilcoxon sign-ranked test is shown. **c**, Pie charts displaying the mean percentage of single CTCs (grey), CTC clusters (green) and CTC–neutrophil clusters (gold) in mice injected with 4T1-pLOC, 4T1-*TLE1* WT, 4T1-*TLE1* G1787A or 4T1-*TLE1* G1509C cells (top). The number of independent biological replicates (n) is shown for each condition. The plots show the mean fold change of CTC ratios (bottom). $n = 3$; error bars, s.e.m.; P values by two-sided Student's t-test are shown.

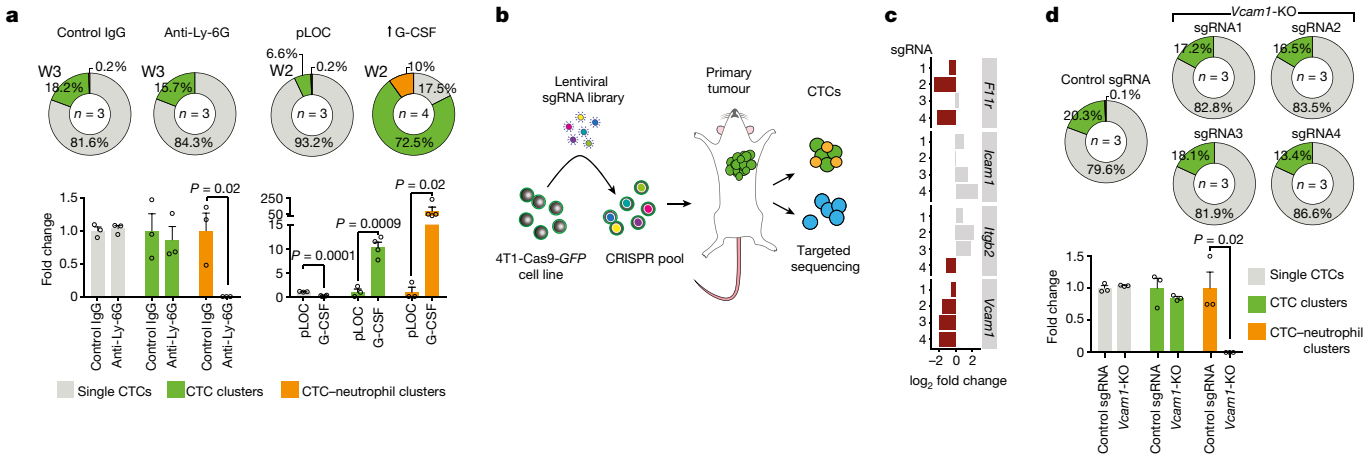


Fig. 4 | Identification of vulnerabilities of CTC–neutrophil clusters. **a**, Pie charts displaying the mean percentage of single CTCs (grey), CTC clusters (green) and CTC–neutrophil clusters (gold) in NSG-4T1-*GFP* mice treated with anti-Ly-6G antibodies or G-CSF overexpression (top). W, weeks upon tumour development. The number of independent biological replicates (*n*) is shown for each condition. The plots show the mean fold change of CTC ratios from NSG-4T1-*GFP* mice treated with anti-Ly-6G antibodies or G-CSF overexpression (bottom). Error bars, s.e.m.; *P* values by two-sided Student's *t*-test are shown. **b**, Schematic of the experimental design. 4T1-Cas9-*GFP* cells are transduced with a pool of sgRNAs targeting cell–cell junctions and injected in NSG mice. Upon tumour development, spontaneously generated CTC–neutrophil clusters, CTC clusters and single CTCs are sequenced to identify sgRNA dropouts. **c**, Bar plot showing the fold change (log_2) of individual sgRNAs (numbered 1 to 4) found in CTCs from CTC–neutrophil clusters versus CTCs alone. sgRNAs with a representation that was reduced in CTCs from CTC–neutrophil clusters are shown in dark red ($n = 3$). **d**, Pie charts displaying the mean percentage of single CTCs (grey), CTC clusters (green) and CTC–neutrophil clusters (gold) in NSG mice carrying 4T1-Cas9-*GFP* tumours expressing a control sgRNA or individual sgRNAs targeting *Vcam1* (*Vcam1*-KO) (top). The number of independent biological replicates (*n*) is shown for each condition. The plot shows the mean fold change of CTC ratios from mice carrying 4T1-Cas9-*GFP* tumours expressing a control sgRNA or individual sgRNAs targeting *Vcam1* (bottom). Error bars, s.e.m., *P* value by two-sided Student's *t*-test is shown.

treatment with neutralizing antibodies against Ly-6G (anti-Ly-6G) or, conversely, we stably overexpressed G-CSF to stimulate the production and recruitment of neutrophils to the tumour site (Extended Data Fig. 9a). As expected, treatment with anti-Ly-6G reduced neutrophil infiltration to the primary tumour site, whereas G-CSF augmented it without altering primary tumour size (Extended Data Fig. 9b, c). However, anti-Ly-6G-treated mice completely lacked CTC–neutrophil clusters from the circulation and displayed a delayed CTC shedding rate compared to control mice, whereas overexpression of G-CSF led to earlier CTC release and substantially increased the proportion of CTC–neutrophil clusters (more than 88-fold) (Fig. 4a, Extended Data Fig. 9d–f). Consequently, neutrophil depletion or augmentation results in a delayed or accelerated metastasis development, respectively, mirrored by a shorter or longer overall survival of treated mice (Extended Data Fig. 9g, h). By contrast, neutrophil depletion with anti-Ly-6G is not effective when cancer cells are administered directly through intravenous injection of pre-treated mice (Extended Data Fig. 9i–l). Of note, in our cohort, G-CSF treatment of patients with breast cancer occurred more often in those patients that were positive for CTCs, including CTC–WBC clusters (Extended Data Fig. 9m). Thus, overall neutrophil abundance affects the likelihood that a tumour has to spontaneously shed CTC–neutrophil clusters.

We next sought to identify actionable vulnerabilities of CTC–neutrophil clusters without targeting the entire neutrophil population. To this end, we investigated cell–cell junction pairs expressed by CTC–neutrophil clusters and possibly mediating their heterotypic cell binding (Extended Data Fig. 10a, b). We engineered a CRISPR–Cas9-based loss-of-function mini-pool screen in vivo, whereby a pool of cells carrying individual knockouts for *F11r*, *Icam1*, *Itgb2* and *Vcam1* are injected in the mammary gland of recipient mice, followed by CTC-targeted barcode sequencing to reveal selective single-guide RNA (sgRNA) dropouts, highlighting genes the knockout of which does not allow the formation of CTC–neutrophil clusters (Fig. 4b). Of note, we observed no differences in primary tumour growth and no selective sgRNA dropouts in primary tumour cells (Extended Data Fig. 10c, d), suggesting that knockout of *F11r*, *Icam1*, *Itgb2* or *Vcam1* does not affect proliferation in the primary tumour. However, we found that four out of four *Vcam1* sgRNAs selectively dropped out in CTCs from CTC–neutrophil clusters but were still present in CTCs alone (Fig. 4c), highlighting a possible VCAM1 requirement for CTC–neutrophil cluster formation. We further validated this finding using individual sgRNAs (Fig. 4d, Extended Data Fig. 10e). Thus, VCAM1 functionally mediates the interaction between CTCs and neutrophils, and its inhibition prevents the formation of CTC–neutrophil clusters.

Altogether, our data provide insights into the processes that define the interaction between cancer cells and immune cells during blood-borne dissemination. We propose a model in which neutrophils directly interact with CTCs to support cell cycle progression in circulation and to accelerate metastasis seeding. This mechanism of metastatic spread and the possibility that CTC–neutrophil clusters may be targeted therapeutically provides an opportunity to reduce the spread of breast cancer.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at https://doi.org/10.1038/s41586-019-0915-y.

Received: 25 January 2018; Accepted: 21 December 2018;
Published online: 06 February 2019

1. Khalil, D. N., Smith, E. L., Brentjens, R. J. & Wolchok, J. D. The future of cancer treatment: immunomodulation, CARs and combination immunotherapy. *Nat. Rev. Clin. Oncol.* **13**, 273–290 (2016).
2. Mohme, M., Riethdorf, S. & Pantel, K. Circulating and disseminated tumour cells – mechanisms of immune surveillance and escape. *Nat. Rev. Clin. Oncol.* **14**, 155–167 (2017).
3. Lambert, A. W., Pattabiraman, D. R. & Weinberg, R. A. Emerging biological principles of metastasis. *Cell* **168**, 670–691 (2017).
4. Aceto, N. et al. Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis. *Cell* **158**, 1110–1122 (2014).
5. Cheung, K. J. et al. Polyclonal breast cancer metastases arise from collective dissemination of keratin 14-expressing tumor cell clusters. *Proc. Natl Acad. Sci. USA* **113**, E854–E863 (2016).
6. Pantel, K. & Speicher, M. R. The biology of circulating tumor cells. *Oncogene* **35**, 1216–1224 (2016).
7. Aceto, N., Toner, M., Maheswaran, S. & Haber, D. A. En route to metastasis: circulating tumor cell clusters and epithelial-to-mesenchymal transition. *Trends Cancer* **1**, 44–52 (2015).

8. Stott, S. L. et al. Isolation of circulating tumor cells using a microvortex-generating herringbone-chip. *Proc. Natl Acad. Sci. USA* **107**, 18392–18397 (2010).
9. Xu, L. et al. Optimization and evaluation of a novel size based circulating tumor cell isolation system. *PLoS ONE* **10**, e0138032 (2015).
10. Li, H. et al. Reference component analysis of single-cell transcriptomes elucidates cellular heterogeneity in human colorectal tumors. *Nat. Genet.* **49**, 708–718 (2017).
11. Fridlender, Z. G. et al. Polarization of tumor-associated neutrophil phenotype by TGF-beta: 'N1' versus 'N2' TAN. *Cancer Cell* **16**, 183–194 (2009).
12. Cristofanilli, M. et al. Circulating tumor cells, disease progression, and survival in metastatic breast cancer. *N. Engl. J. Med.* **351**, 781–791 (2004).
13. Labelle, M., Begum, S. & Hynes, R. O. Direct signaling between platelets and cancer cells induces an epithelial-mesenchymal-like transition and promotes metastasis. *Cancer Cell* **20**, 576–590 (2011).
14. Rothstein, G. et al. Stimulation of neutrophil production in CSF-1-responsive clones. *Blood* **72**, 898–902 (1988).
15. He, J. Q. et al. Association of genetic variations in the *CSF2* and *CSF3* genes with lung function in smoking-induced COPD. *Eur. Respir. J.* **32**, 25–34 (2008).
16. Verri, W. A., Jr et al. IL-15 mediates antigen-induced neutrophil migration by triggering IL-18 production. *Eur. J. Immunol.* **37**, 3373–3380 (2007).
17. Kacinski, B. M. et al. FMS (CSF-1 receptor) and CSF-1 transcripts and protein are expressed by human breast carcinomas in vivo and in vitro. *Oncogene* **6**, 941–952 (1991).
18. McCracken, J. M. & Allen, L. A. Regulation of human neutrophil apoptosis and lifespan in health and disease. *J. Cell Death* **7**, 15–23 (2014).
19. Canli, O. et al. Myeloid cell-derived reactive oxygen species induce epithelial mutagenesis. *Cancer Cell* **32**, 869–883 (2017).
20. Alexandrov, L. B. et al. Signatures of mutational processes in human cancer. *Nature* **500**, 415–421 (2013).
21. Wellenstein, M. D. & de Visser, K. E. Cancer-cell-intrinsic mechanisms shaping the tumor immune landscape. *Immunity* **48**, 399–416 (2018).
22. Ramasamy, S. et al. *Tle1* tumor suppressor negatively regulates inflammation in vivo and modulates NF-κB inflammatory pathway. *Proc. Natl Acad. Sci. USA* **113**, 1871–1876 (2016).

Acknowledgements We thank all patients that donated blood for our study, as well as all involved clinicians and study nurses; J. Massagué (Memorial Sloan Kettering Cancer Center), D. Haber and S. Maheswaran (Massachusetts General Hospital and Harvard Medical School) for donating cell lines; G. Christofori for MMTV-PyMT mice and comments on the manuscript, and all members of the Aceto laboratory for feedback and discussions; K. Eschbach and E. Burcklen from the Genomics Facility Basel (D-BSSE of the ETH Zürich) for generating sequencing libraries and performing next-generation sequencing; S. Arnold (D-BSSE of the ETH Zürich) and S. Münst Soysal (University Hospital Basel) for support with sample acquisition and processing; and T. Ryser (Aceto laboratory, University of Basel) for help with CRISPR–Cas9-related experiments. Calculations were performed at sciCORE (http://scicore.unibas.ch/) scientific computing center of the University of Basel. Research in the Aceto laboratory is supported by the European Research Council, the Swiss National Science Foundation, the Swiss Cancer League, the Basel Cancer League, the two Cantons of Basel through the ETH Zürich and the University of Basel.

Reviewer information *Nature* thanks K. Pantel and the other anonymous reviewer(s) for their contribution to the peer review of this work.

Author contributions B.M.S. and N.A. designed the study, performed the experiments and wrote the manuscript. F.C.-G. performed the computational analysis. S.G., I.K., C.D. and R.S. processed blood samples, mouse tissues and performed immunofluorescence staining. M.C.S. performed CRISPR–Cas9-related experiments. M.V., J.L., C.K., V.H.-S., C.R. and W.P.W. provided patient samples and clinical input throughout the project. C.B. generated sequencing data. J.S. and N.B. provided input during computational methods development and data analysis. All authors have read, commented and approved the manuscript in its final form.

Competing interests N.A. and B.M.S. are listed as inventors in patent applications that are related to CTC clusters and CTC–neutrophil clusters.

Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41586-019-0915-y.
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-019-0915-y.
Reprints and permissions information is available at http://www.nature.com/reprints.
Correspondence and requests for materials should be addressed to N.A.
Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Breast Cancer Research

RESEARCH ARTICLE Open Access

Denosumab treatment is associated with the absence of circulating tumor cells in patients with breast cancer

Marcus Vetter[1,2†], Julia Landin[2†], Barbara Maria Szczerba[3†], Francesc Castro-Giner[3,4†], Sofia Gkountela[3], Cinzia Donato[3], Ilona Krol[3], Ramona Scherrer[3], Catharina Balmelli[2], Alexandra Malinovska[2], Alfred Zippelius[2], Christian Kurzeder[1,5], Viola Heinzelmann-Schwarz[1], Walter Paul Weber[5], Christoph Rochlitz[2] and Nicola Aceto[3*]

Abstract

Background: The presence of circulating tumor cells (CTCs) in patients with breast cancer correlates to a bad prognosis. Yet, CTCs are detectable in only a minority of patients with progressive breast cancer, and factors that influence the abundance of CTCs remain elusive.

Methods: We conducted CTC isolation and enumeration in a selected group of 73 consecutive patients characterized by progressive invasive breast cancer, high tumor load and treatment discontinuation at the time of CTC isolation. CTCs were quantified with the Parsortix microfluidic device. Clinicopathological variables, blood counts at the time of CTC isolation and detailed treatment history prior to blood sampling were evaluated for each patient.

Results: Among 73 patients, we detected at least one CTC per 7.5 ml of blood in 34 (46%). Of these, 22 (65%) had single CTCs only, whereas 12 (35%) featured both single CTCs and CTC clusters. Treatment with the monoclonal antibody denosumab correlated with the absence of CTCs, both when considering all patients and when considering only those with bone metastasis. We also found that low red blood cell count was associated with the presence of CTCs, whereas high CA 15-3 tumor marker, high mean corpuscular volume, high white blood cell count and high mean platelet volume associated specifically with CTC clusters.

Conclusions: In addition to blood count correlatives to single and clustered CTCs, we found that denosumab treatment associates with most patients lacking CTCs from their peripheral circulation. Prospective studies will be needed to validate the involvement of denosumab in the prevention of CTC generation.

Keywords: Circulating tumor cells, Circulating tumor cell clusters, Denosumab, Breast cancer, Metastasis

Background

Circulating tumor cells (CTCs) are derivatives of solid tumor lesions that detach from the tumor and enter the bloodstream [1]. In patients with breast cancer, CTCs have been shown to be predictive of a shorter disease-free survival and overall survival [2, 3], with a worse prognosis in patients who present with a count of at least five CTCs per 7.5 ml of blood [2, 3]. Generally, high CTC counts have been associated with a poor prognosis in multiple settings, including those patients that are newly diagnosed with metastatic breast cancer and about to start a therapy [3, 4]. At the morphological level, breast CTCs occur in the blood of patients as single CTCs or as CTC clusters, with the latter being associated with a shorter metastasis-free survival than in patients in whom only single CTCs are found [5].

Although the association between CTCs and bad prognosis is well established in breast cancer, CTCs are detectable only in a subset (~ 20–40%) of patients [2, 5]. To date, no parameters have been found that could explain CTC abundance in patients, leading to difficulties in enabling patient stratification prior to CTC-related

* Correspondence: nicola.aceto@unibas.ch
†Marcus Vetter, Julia Landin, Barbara Maria Szczerba and Francesc Castro-Giner contributed equally to this work.
[3]Department of Biomedicine, Cancer Metastasis Laboratory, University of Basel and University Hospital Basel, Mattenstrasse 28, CH-4058 Basel, Switzerland
Full list of author information is available at the end of the article

investigations [6, 7], as well as limiting our understanding of those factors that may influence the spread of cancer.

In this study, we aimed to investigate a number of clinicopathological variables, blood counts at the time of CTC isolation and detailed treatment history prior to blood sampling in a cohort of 73 consecutive patients with invasive breast cancer characterized by progressive disease, high tumor load and treatment discontinuation (or without any pretreatment) at the time of CTC isolation, before the next line of therapy. Additionally, we not only investigated parameters that are associated with the presence of CTCs but also specifically interrogated our datasets to identify features that are associated with CTC clusters. The rationale of our study was therefore to identify, in an unbiased manner (i.e., not driven by preexisting hypotheses), clinical parameters that correlate with CTC presence in patients with progressive breast cancer.

Methods

Patient selection

Seventy-three consecutive patients with invasive breast cancer, progressive disease, high tumor load, treatment discontinuation at the time of CTC isolation (before the next line of therapy) and no preselection for breast cancer subtype or specific metastatic sites were enrolled in the study. Eligible patients were > 18 years old with any menopausal status and had an Eastern Cooperative Oncology Group performance status of 0–3. Disease had to be measurable by Response Evaluation Criteria in Solid Tumors (RECIST) version 1.1 or nonmeasurable bone-only disease. Tumor load was defined by either the size of the primary tumor or the number and size of metastatic lymph nodes or distant sites, and patients with higher tumor load were prioritized. All subjects donated 7.5–15 ml of blood in ethylenediaminetetraacetic acid (EDTA) vacutainers at least once, and each signed an informed consent before joining the study. The study was performed under the protocols EKNZ BASEC 2016-00067 and EK 321/10, which received ethical and institutional review board approvals before study initiation (Ethics Committee northwest/central Switzerland [EKNZ]). This study was performed in compliance with the Declaration of Helsinki.

CTC isolation and enumeration strategy

Patient-derived CTCs were captured on the microfluidic Parsortix Cell Separation Cassette (GEN3D6.5; ANGLE, Guildford, UK) within 1 h of blood draw, directly from unmanipulated blood samples. Next, in-cassette staining was performed with an antibody cocktail comprising antibodies against epithelial cell adhesion molecule (EpCAM)-Alexa Fluor 488 (AF488) (#CST5198; Cell Signaling Technology, Danvers, MA, USA), human epidermal growth factor receptor 2 (HER2)-AF488 (#324410; BioLegend, San Diego, CA, USA), epidermal growth factor receptor (EGFR)-fluorescein isothiocyanate (FITC) (#GTX11400; GeneTex, Irvine, CA, USA) and CD45-BV605 (#304042; BioLegend). CTCs were characterized as AF488/FITC-positive and BV605-negative and enumerated manually by two independent operators under a fluorescence microscope at 20× magnification.

Clinical parameter assessment

Primary tumor samples were collected at the initial diagnosis, and IHC was performed for estrogen receptor (ER), progesterone receptor (PR), HER2 and Ki-67. If the patient had primary metastatic disease, a biopsy from the metastatic site was obtained when possible, including marker assessment: ER, PR and HER2. Histopathological diagnosis was conducted by two independent pathologists from the breast cancer unit at the University Hospital Basel. All patients were treated at the Breast Cancer Unit University Hospital Basel according to local standard operating procedures and National Comprehensive Cancer Network and European Society for Medical Oncology guidelines by senior breast oncologists. If a patient had a progression within new distant sites, a new biopsy from that site was taken, when possible, to determine ER, PR and HER2. Patients under systemic treatment had tumor assessment at least every 12 weeks with computed tomographic scans or earlier if tumor progression was anticipated. CTC collection was performed at progression and prior to the next line of therapy or before any treatment was conducted. The patients' data was retrieved by detailed retrospective chart review. Data collection included demographics and disease-specific and treatment-specific data including age, gender, primary stage, histologic subtype, ER/PR/HER2 status, grading, Ki-67, date of primary diagnosis and relapse, type of relapse (localized, metastatic), site of distant disease, bone-modifying agents (bisphosphonates, denosumab), palliative irradiation, and type of systemic treatment, including time on treatment and time to next subsequent treatment. Data was correlated with CTC counts.

Blood parameter assessment

Complete blood counts were measured with the ADVIA 120 Hematology Analyzer (Siemens Healthcare Diagnostics, Tarrytown, NY, USA) using Multispecies version 5.9.0-MS software (Bayer Diagnostics, Tarrytown, NY, USA). Blood samples were taken before each new therapy cycle or at least every month, including cancer antigen 15-3 (CA 15-3), alkaline phosphatase, Ca^{2+}, C-reactive protein, lactate dehydrogenase, red blood cells (RBC), hemoglobin, hematocrit, mean corpuscular volume (MCV), mean corpuscular hemoglobin concentration, white blood cells

(WBC), neutrophils, lymphocytes, monocytes, eosinophils, basophils, large unstained cells, platelets and mean platelet volume (MPV). In the vast majority of cases, blood samples were taken simultaneously with the CTC sample or within 7 days after CTCs were taken. Eight of 73 patients had only partial data available, whereas no blood counts were reported at the time of CTC detection for nine of 73 patients.

Statistical analysis

We first screened our data to exclude variables and patients with high content of missing information, as well as observations with implausible values. Cancer therapies were simplified into three main nonexclusive categories (targeted therapy, chemotherapy and hormone therapy) (Additional file 1: Table S1). Some patients had undergone multiple lines of therapy. For this reason, we assessed the effects of accumulated therapies and the therapy at CTC evaluation separately.

We investigated the association between the different variables of interest and the presence of CTCs using Fisher's exact test for categorical variables, two-sided Wilcoxon rank-sum test for continuous variables (e.g., complete blood counts) and Kruskal-Wallis test for ordinal variables with more than two levels (e.g., stage at diagnosis). A list of the statistical tests used for each variable can be found in Additional file 2: Table S2. For each test, we present the nominal P value. An estimate and 95% CI are also provided for continuous and two-level categorical variables. The estimate corresponds to the OR in Fisher's exact test and to the estimated median of the difference between samples from both groups in the Wilcoxon rank-sum test. To account for potential confounding variables, logistic regression analysis was conducted, adjusting by age at primary diagnosis, tumor stage at diagnosis, tumor grade and histologic subtype. Adjusted P values were calculated following the Benjamini-Hochberg method, combining all tests performed in this work. Associations with an adjusted P value ≤ 0.05 are highlighted in the text. We conducted the data wrangling and statistical analysis in R (version 3.4.0; R Foundation for Statistical Computing, Vienna, Austria).

Results

Patient characteristics

Given previously reported correlations between number of CTCs and tumor load [8], as well as the findings that CTC counts predict poor prognosis in breast cancer [3, 4, 9], we focused on a group of 73 consecutive patients with invasive breast cancer with the following characteristics: high tumor load, detailed treatment history available, progressive disease associated with treatment discontinuation at the time of CTC isolation (before the next line of therapy), and availability of comprehensive blood counts performed at CTC collection. Selected patients ranged from 36 to 85 years of age and carried either invasive ductal, lobular or inflammatory carcinoma, with a broad expression range of ER, PR, HER2 and Ki-67 protein levels, as well as tumor grade varying from 1 to 3. Detailed characteristics of patients, therapies and statistical tests used for the analysis are listed in Additional file 1: Table S1, Additional file 2: Table S2, and Additional file 3: Table S3.

CTC isolation and enumeration

Blood samples were drawn in EDTA vacutainers and processed directly with the Parsortix microfluidic device [10], with a dedicated protocol enabling the isolation of > 99% of breast CTCs from unlabeled blood samples (Additional file 4: Figure S1A). Upon enrichment, CTCs were stained with an antibody cocktail against EpCAM, EGFR and HER2 and counterstained for the WBC marker CD45 (Additional file 4: Figure S1A and B). With this approach, we detected at least one CTC per 7.5 ml of peripheral blood in 34 (46.6%) patients. Among these, we observed that 22 (64.7%) patients were characterized by the presence of single CTCs and that 12 (35.3%) patients had both single CTCs and CTC clusters (Additional file 4: Figure S1C).

Features of patients with CTCs

We investigated a number of clinicopathological variables to identify features associated with patients in whom either single CTCs or CTC clusters were found, compared with patients with no detectable CTCs. We first observed that previous treatments with targeted therapy (including but not limited to hormonal, anti-HER2, anti-CDK4/6 treatments), chemotherapy or radiotherapy did not correlate with the presence of CTCs (Additional file 5: Table S4). Yet, we found that treatment with the anti-bone resorption antibody denosumab (received by 21 of 73 patients) was associated with the absence of CTCs (OR, 0.25; 95% CI, 0.06–0.86; $P = 0.019$). Namely, the prevalence of CTCs was 14.7% (5 of 21) among patients treated with denosumab and 55.8% (29 of 52) among nontreated patients (Table 1). Further, when considering only those patients in whom CTCs were detected, the average CTC number was 9.8 for patients treated with denosumab ($n = 5$) versus 24.79 for nontreated patients ($n = 29$). Despite their role as anti-bone resorption agents, the same association was not seen for bisphosphonates ($P = 0.784$). Importantly, anti-bone resorption treatment with either denosumab or bisphosphonates was decided on the basis of treatment initiation date (denosumab was approved in Switzerland in December 2011 and given as the preferred treatment option to eligible patients after that date), whereas patients who started receiving bisphosphonates (i.e., prior to December 2011) continued receiving bisphosphonates unless major side effects

Table 1 Clinical features of patients with circulating tumor cells

	No CTC ($n = 39$)	CTC ($n = 34$)	*P* value	Estimate (95% CI)
Age at primary diagnosis, years, mean (SD)	58.38 (11.85)	55.1 (11.04)	0.444	− 2.76 (− 8.49, 2.84)
Age at first CTC evaluation, years, mean (SD)	63.53 (11.69)	59.58 (10.7)	0.163	− 4.11 (− 9.27, 1.63)
Stage at diagnosis, *n* (%)			0.679	–
I	4 (10.53%)	5 (14.71%)		
IA	1 (2.63%)	0 (0%)		
II	5 (13.16%)	4 (11.76%)		
IIA	1 (2.63%)	4 (11.76%)		
III	11 (28.95%)	7 (20.59%)		
IIIA	2 (5.26%)	0 (0%)		
IIIC	0 (0%)	2 (5.88%)		
IV	14 (36.84%)	11 (32.35%)		
Lymphocyte node involvement, *n* (%)			0.881	–
N0	11 (31.43%)	8 (25.81%)		
N1	11 (31.43%)	14 (45.16%)		
N2	6 (17.14%)	0 (0%)		
N3	6 (17.14%)	9 (29.03%)		
Histologic subtype, *n* (%)			0.964	–
Invasive lobular	6 (15.38%)	4 (11.76%)		
Invasive ductal	31 (79.49%)	29 (85.29%)		
Inflammatory invasive lobular	1 (2.56%)	1 (2.94%)		
Inflammatory	2 (5.13%)	2 (5.88%)		
% of ER^+ cells, mean (SD)	65.82 (42.48)	59.79 (40.93)	0.386	0 (− 10, 0)
% of PR^+ cells, mean (SD)	38.97 (40.04)	26.35 (33.78)	0.171	− 4 (− 20, 0)
% of Ki-67^+ cells, mean (SD)	27.08 (17.13)	31.59 (21.55)	0.514	5 (− 5, 10)
$HER2^+$	7 (17.95%)	7 (22.58%)	0.766	1.33 (0.35–5.11)
Triple-negative	4 (10.81%)	3 (9.68%)	1.000	0.89 (0.12–5.73)
Tumor grade			0.985	–
1	4 (10.53%)	4 (12.12%)		
2	17 (44.74%)	14 (42.42%)		
3	17 (44.74%)	15 (45.45%)		
Bisphosphonates	9 (23.68%)	7 (20.59%)	0.784	0.84 (0.23–2.94)
Denosumab	16 (41.03%)	5 (14.71%)	0.019	0.25 (0.06–0.86)
Radiotherapy	22 (56.41%)	13 (38.24%)	0.160	0.48 (0.17–1.35)
Relapse				
Any	31 (79.49%)	25 (73.53%)	0.589	0.72 (0.21–2.45)
Local	3 (7.69%)	4 (11.76%)	0.698	1.59 (0.25–11.72)
Metastasis	26 (66.67%)	19 (55.88%)	0.470	0.64 (0.22–1.82)
Days between primary diagnosis and relapse, mean (SD)	1954.08 (2042.25)	1893.48 (1853.86)	0.966	− 9.37 (− 1000, 756)
Established metastatic disease at CTC evaluation	35 (89.74%)	30 (88.24%)	1.000	0.86 (0.15–5.04)
Number of metastatic sites, mean (SD)	2.09 (1.01)	1.9 (0.94)	0.473	0 (− 1, 0)
Metastasis site, *n* (%)				
Bone	27 (69.23%)	17 (51.52%)	0.150	0.45 (0.15–1.28)
Liver	10 (25.64%)	12 (36.36%)	0.447	1.57 (0.51–4.9)
Lymph node	9 (23.08%)	10 (30.3%)	0.599	1.38 (0.43–4.54)

Table 1 Clinical features of patients with circulating tumor cells *(Continued)*

	No CTC ($n = 39$)	CTC ($n = 34$)	*P* value	Estimate (95% CI)
Pleural	7 (17.95%)	2 (6.06%)	0.162	0.29 (0.03–1.68)
Peritoneal	3 (7.69%)	4 (12.12%)	0.698	1.59 (0.25–11.72)
Lung	4 (10.26%)	4 (12.12%)	1.000	1.16 (0.2–6.83)
Skin	3 (7.69%)	0 (0%)	0.243	0 (0–2.74)
Brain	2 (5.13%)	2 (6.06%)	1.000	1.15 (0.08–16.76)
Uterus	1 (2.56%)	1 (3.03%)	1.000	1.15 (0.01–92.67)
Muscular	1 (2.56%)	2 (6.06%)	0.595	2.35 (0.12–143.61)

Abbreviations: ER Estrogen receptor, *HER2* Human epidermal growth factor receptor 2, *PR* Progesterone receptor
The table shows clinical features of patients with and without circulating tumor cells (CTCs)

occurred (Additional file 6: Table S5). When we restricted the analysis to the 44 patients with bone metastasis, denosumab was administered to 20 of 44 of them, and it also correlated with a reduction in CTC numbers compared with the remaining 24 patients with bone metastasis but no denosumab treatment (OR, 0.22; 95% CI, 0.04–0.96; $P = 0.03$) (Table 2). These results were confirmed using logistic regression adjusting by age at primary diagnosis, tumor stage at diagnosis, tumor grade and histologic subtype (OR, 0.25; 95% CI, 0.06–0.82; $P = 0.03$). When comparing clinicopathological variables in patients who were treated or not with denosumab, as expected, we observed a correlation with bone metastasis (OR, 22.53; 95% CI, 3.14–995.64; $P = 5.6e\text{-}05$; adjusted $P = 0.01$) (Additional file 7: Table S6), but no effect on progression-free survival was seen (Additional file 8: Figure S2). Together, our data show that denosumab treatment is associated with a marked reduction of CTC counts in patients with breast cancer.

Features of patients with CTC clusters

We further asked whether any clinicopathological variables might be associated specifically with the presence of CTC clusters, compared with patients in whom CTC clusters were not found (i.e., having either single CTCs or no CTCs). We found that both younger age at primary diagnosis and younger age at first CTC evaluation were associated with the presence of CTC clusters (Table 3). Particularly, we observed average ages at primary diagnosis of 50.63 years (SD, 12.60) for patients with CTC clusters and 58.08 years (SD, 10.99) for patients with no CTC clusters ($P = 0.033$), as well as average ages at first CTC evaluation of 54.87 years (SD, 12.14) for patients with CTC clusters and 63.03 years (SD, 10.77) for patients with no CTC clusters ($P = 0.025$). We also observed that although HER2 was expressed in 22% (13 of 61) of patients with no CTC clusters (7 of 39) (i.e., 17.95% of patients with no CTCs and 6 of 22 [30%] of patients with single CTCs only), it was expressed in only 1 patient (9.09%) with CTC clusters (OR, 0.36; 95% CI, 0.01–2.97; $P = 0.44$), even though the relationship between HER2 negativity and CTC clusters did not reach statistical significance (Table 3).

When considering only patients with CTCs and comparing those with CTC clusters versus those with single CTCs, we found that also in this context younger age at primary diagnosis ($P = 0.044$) and younger age at first CTC evaluation ($P = 0.058$) were associated with the presence of CTC clusters (Additional file 9: Table S7).

Blood parameters associated with CTCs

In addition to investigating the clinical parameters summarized above, for each patient, we also evaluated comprehensive blood counts performed at CTC collection. We first asked whether blood-related parameters were associated with the presence of CTCs (either single or clustered), compared with patients in whom CTCs were not detected. We observed that patients with detectable CTCs had a lower RBC count (OR, − 0.42; 95% CI, − 0.8 - -0.08; $P = 0.019$) than patients with no CTCs (Table 4).

Blood parameters associated with CTC clusters

We then asked whether specific blood-related parameters could be associated with the presence of CTC clusters, compared with patients with no CTC clusters (i.e., having either no CTCs or single CTCs only). In this

Table 2 Circulating tumor cells detection according to denosumab treatment and bone metastasis

	Number of samples	No CTCs	CTCs	*P* value	Estimate (95% CI)	*P* value	Estimate (95% CI)
Reference	28	12 (43%)	16 (57%)	Reference		–	–
Bone metastasis	24	11 (46%)	13 (54%)	1.000	0.89 (0.26–3.05)	Reference	
Bone metastasis and denosumab	20	16 (80%)	4 (20%)	0.017	0.19 (0.04–0.81)	0.030	0.22 (0.04–0.96)

The table shows the number of patients with and without circulating tumor cells (CTCs) among individuals with bone metastasis who were treated or not with denosumab

Table 3 Clinical features of patients with circulating tumor cell clusters

	No CTC clusters ($n = 61$)	CTC clusters ($n = 12$)	*P* value	Estimate (95% CI)
Age at primary diagnosis, years, mean (SD)	58.08 (10.99)	50.63 (12.6)	0.033	− 8.26 (− 15.3, − 0.44)
Age at first CTC evaluation, years, mean (SD)	63.03 (10.77)	54.87 (12.14)	0.025	− 8.3 (− 16.06, − 1.04)
Stage at diagnosis, *n* (%)			0.726	–
I	7 (11.67%)	2 (16.67%)		
IA	1 (1.67%)	0 (0%)		
II	7 (11.67%)	2 (16.67%)		
IIA	4 (6.67%)	1 (8.33%)		
III	16 (26.67%)	2 (16.67%)		
IIIA	2 (3.33%)	0 (0%)		
IIIC	1 (1.67%)	1 (8.33%)		
IV	21 (35%)	4 (33.33%)		
Lymphocyte node involvement, *n* (%)			0.855	–
N0	15 (27.27%)	4 (36.36%)		
N1	22 (40%)	3 (27.27%)		
N2	6 (10.91%)	0 (0%)		
N3	11 (20%)	4 (36.36%)		
Histologic subtype, *n* (%)			0.679	–
Invasive lobular	9 (14.75%)	1 (8.33%)		
Invasive ductal	49 (80.33%)	11 (91.67%)		
Inflammatory invasive lobular	1 (1.64%)	1 (8.33%)		
Inflammatory	3 (4.92%)	1 (8.33%)		
% of ER^+ cells, mean (SD)	62.34 (41.8)	66.42 (42.1)	0.675	0 (− 10, 20)
% of PR^+ cells, mean (SD)	32.71 (37.79)	34 (37.3)	0.888	0 (− 10, 20)
% of Ki-67^+ cells, mean (SD)	30 (19.65)	23 (16.43)	0.384	− 5 (− 20, − 10)
$HER2^+$, *n* (%)	13 (22.03%)	1 (9.09%)	0.442	0.36 (0.01–2.97)
Triple-negative, *n* (%)	7 (12.28%)	0 (0%)	0.588	0 (0–3.7)
Tumor grade, *n* (%)			0.093	–
1	5 (8.33%)	3 (27.27%)		
2	26 (43.33%)	5 (45.45%)		
3	29 (48.33%)	3 (27.27%)		
Bisphosphonates, *n* (%)	14 (23.33%)	2 (16.67%)	1.000	0.66 (0.06–3.68)
Denosumab, *n* (%)	19 (31.15%)	2 (16.67%)	0.489	0.45 (0.04–2.41)
Radiotherapy, *n* (%)	30 (49.18%)	5 (41.67%)	0.756	0.74 (0.17–3.06)
Relapse, *n* (%)				
Any	47 (77.05%)	9 (75%)	1.000	0.9 (0.19–5.83)
Local	4 (6.56%)	3 (25%)	0.082	4.61 (0.58–32.62)
Metastasis	40 (65.57%)	5 (41.67%)	0.193	0.38 (0.08–1.59)
Days between primary diagnosis and relapse, mean (SD)	1969.49 (2003.96)	1636.67 (1538.48)	0.633	− 236.86 (− 1643, 1203)
Established metastatic disease at CTC evaluation, *n* (%)	54 (88.52%)	11 (91.67%)	1.000	1.42 (0.15–70.01)
Number of metastatic sites, mean (SD)	1.96 (0.98)	2.18 (0.98)	0.452	0 (0–1)
Metastasis site, *n* (%)				
Bone	37 (61.67%)	7 (58.33%)	1.000	0.91 (0.22–4.08)
Liver	19 (31.67%)	3 (25%)	1.000	0.74 (0.12–3.42)
Lymph node	15 (25%)	4 (33.33%)	0.497	1.52 (0.29–6.74)

Table 3 Clinical features of patients with circulating tumor cell clusters *(Continued)*

	No CTC clusters ($n = 61$)	CTC clusters ($n = 12$)	*P* value	Estimate (95% CI)
Pleural	9 (15%)	0 (0%)	0.339	0 (0–2.57)
Peritoneal	5 (8.33%)	2 (16.67%)	0.323	2.21 (0.19–16.05)
Lung	7 (11.67%)	1 (8.33%)	1.000	0.7 (0.01–6.47)
Skin	3 (5%)	0 (0%)	1.000	0 (0–12.81)
Brain	3 (5%)	1 (8.33%)	0.521	1.74 (0.03–24.14)
Uterus	1 (1.67%)	1 (8.33%)	0.304	5.27 (0.06–433.34)
Muscular	2 (3.33%)	1 (8.33%)	0.421	2.63 (0.04–54.78)

Abbreviations: ER Estrogen receptor, *HER2* Human epidermal growth factor receptor 2, *PR* Progesterone receptor
The table shows clinical features of patients with and without circulating tumor cell clusters (CTC clusters)

case, we found that patients with CTC clusters have 14-fold higher levels of the CA 15-3 tumor marker ($P = 0.021$), higher MCV ($P = 0.033$), higher WBC ($P = 0.03$) and higher MPV ($P = 0.032$) than patients in whom CTC clusters were not found (Table 5). We also restricted this analysis to patients with CTCs and compared patients with CTC clusters with patients with only single CTCs. In this setting, we further confirmed that patients with CTC clusters have 38-fold higher CA 15-3 tumor antigen ($P = 0.0089$), as well as nearly twofold higher total WBC counts ($P = 0.0045$) and higher neutrophil counts ($P = 0.03$) (Additional file 10: Table S8).

Discussion

In a selected cohort of 73 patients with progressive invasive breast cancer, we provide a detailed description of a number of clinicopathological parameters and blood counts at the time of CTC isolation that correlate with the presence of single CTCs and CTC clusters. Interestingly, we observed that treatment with the monoclonal

Table 4 Complete blood counts in patients with circulating tumor cells

	No CTC ($n = 39$)	CTC ($n = 34$)	*P* value	Estimate (95% CI)
CA 15-3, mean (SD)	223.71 (384.68)	1084.15 (4136.87)	0.658	6.7 (– 19.2, 87.6)
Alkaline phosphatase, mean (SD)	105.47 (103.98)	198.15 (365.58)	0.401	6 (– 12, 27)
Calcium (korr), mean (SD)	2.34 (0.15)	2.32 (0.25)	0.145	– 0.06 (– 0.13, 0.02)
CRP, mean (SD)	31.92 (47.56)	26.87 (47.69)	0.982	0 (– 8.8, 3.8)
LDH, mean (SD)	281.61 (118.18)	300.15 (228.57)	0.772	– 5 (– 36, 23)
RBC, 10^{12}/L, mean (SD)	4.37 (0.56)	3.85 (0.77)	0.019	– 0.42 (– 0.8, – 0.08)
HGB, g/L, mean (SD)	130.14 (19.85)	118.15 (24.11)	0.051	– 11 (– 21, 0)
HCT, L/L, mean (SD)	0.38 (0.06)	0.35 (0.06)	0.053	– 0.03 (– 0.06, 0)
MCV, fl, mean (SD)	87.46 (5.68)	89.73 (5.26)	0.227	2 (– 1, 4)
MCH, pg, mean (SD)	29.62 (2.52)	30.66 (1.96)	0.157	0.8 (– 0.3, 1.8)
MCHC, g/L, mean (SD)	339.03 (13.98)	341.19 (12.79)	0.667	2 (– 6, 8)
WBC, 10^9/L, mean (SD)	7.35 (2.13)	7.24 (3.53)	0.334	– 0.6 (– 1.88, 0.81)
Neutrophils, 10^9/L, mean (SD)	5.33 (1.87)	5.12 (2.87)	0.239	– 0.63 (– 1.65, 0.56)
Lymphocytes, 10^9/L, mean (SD)	1.37 (0.67)	1.41 (0.85)	0.941	– 0.02 (– 0.38, 0.38)
Monocytes, 10^9/L, mean (SD)	0.42 (0.12)	0.44 (0.2)	0.843	– 0.01 (– 0.08, 0.07)
Eosinophils, 10^9/L, mean (SD)	0.16 (0.18)	0.17 (0.12)	0.423	0.02 (– 0.03, 0.09)
Basophils, 10^9/L, mean (SD)	0.04 (0.07)	0.04 (0.03)	0.256	0.01 (– 0.01, 0.02)
LUC, 10^9/L, mean (SD)	0.15 (0.22)	0.12 (0.08)	0.912	0 (– 0.03, 0.02)
PLT, 10^9/L, mean (SD)	289.92 (139.86)	249.93 (91.52)	0.277	– 30 (– 84, 24)
MPV, fl, mean (SD)	8.2 (1.51)	8.68 (1.52)	0.109	0 (0–1)

Abbreviations: CA 15-3 Cancer antigen 15-3, *CRP* C-reactive protein, *HCT* Hematocrit, *HGB* Hemoglobin, *LDH* Lactate dehydrogenase, *LUC* Large unstained cells, *MCH* Mean corpuscular hemoglobin, *MCHC* Mean corpuscular hemoglobin concentration, *MCV* Mean corpuscular volume, *MPV* Mean platelet volume, *PLT* Platelets, *RBC* Red blood cells, *WBC* White blood cells
The table shows complete blood counts in patients with and without circulating tumor cells (CTCs).

Table 5 Complete blood counts in patients with circulating tumor cell clusters

	No CTC clusters ($n = 61$)	CTC clusters ($n = 12$)	*P* value	Estimate (95% CI)
CA 15-3, mean (SD)	172.5 (324.45)	2554.6 (6387.64)	0.021	204.16 (9.7–515)
Alkaline phosphatase, mean (SD)	106.4 (104.27)	310.25 (525.57)	0.301	10 (− 12, 74.58)
Calcium (korr), mean (SD)	2.33 (0.14)	2.36 (0.35)	0.698	− 0.02 (− 0.16, 0.14)
CRP, mean (SD)	27.76 (44.47)	37.25 (58.6)	0.279	2.9 (− 4, 21.4)
LDH, mean (SD)	271 (102.59)	363.33 (329.24)	0.463	17 (− 26, 76)
RBC, 10^{12}/L, mean (SD)	4.26 (0.6)	3.66 (0.93)	0.078	− 0.51 (− 1.16, 0.06)
HGB, g/L, mean (SD)	127.52 (20.18)	114.18 (29.19)	0.183	− 12 (− 30, 6)
HCT, L/L, mean (SD)	0.38 (0.05)	0.34 (0.08)	0.159	− 0.03 (− 0.09, 0.01)
MCV, fl, mean (SD)	87.7 (5.58)	91.73 (4.41)	0.033	4 (0–7)
MCH, pg, mean (SD)	29.8 (2.42)	31.24 (1.51)	0.064	1.1 (0–2.3)
MCHC, g/L, mean (SD)	339.8 (13.38)	340.64 (14.25)	0.646	2 (− 10, 10)
WBC, 10^9/L, mean (SD)	6.87 (2.25)	9.38 (3.99)	0.030	2.54 (0.26–4.68)
Neutrophils, 10^9/L, mean (SD)	4.94 (1.87)	6.65 (3.6)	0.177	1.22 (− 0.52, 3.68)
Lymphocytes, 10^9/L, mean (SD)	1.38 (0.69)	1.42 (1)	0.756	− 0.11 (− 0.62, 0.6)
Monocytes, 10^9/L, mean (SD)	0.41 (0.13)	0.5 (0.24)	0.336	0.06 (− 0.08, 0.26)
Eosinophils, 10^9/L, mean (SD)	0.16 (0.16)	0.17 (0.14)	0.974	0 (− 0.07, 0.1)
Basophils, 10^9/L, mean (SD)	0.04 (0.06)	0.04 (0.04)	0.983	0 (− 0.02, 0.02)
LUC, 10^9/L, mean (SD)	0.14 (0.19)	0.12 (0.11)	0.471	− 0.01 (− 0.05, 0.02)
PLT, 10^9/L, mean (SD)	278.02 (125.8)	251.5 (109.99)	0.667	− 15.84 (− 94, 47)
MPV, fl, mean (SD)	8.32 (1.62)	8.82 (0.87)	0.032	1 (0–1)

Abbreviations: CA 15-3 Cancer antigen 15-3, *CRP* C-reactive protein, *HCT* Hematocrit, *HGB* Hemoglobin, *LDH* Lactate dehydrogenase, *LUC* Large unstained cells, *MCH* Mean corpuscular hemoglobin, *MCHC* Mean corpuscular hemoglobin concentration, *MCV* Mean corpuscular volume, *MPV* Mean platelet volume, *PLT* Platelets, *RBC* Red blood cells, *WBC* White blood cells

The table shows complete blood counts in patients with and without circulating tumor cell clusters (CTC clusters)

antibody denosumab in patients with bone metastasis strongly correlated with the absence of CTCs from their peripheral circulation, suggesting a scenario in which the treatment itself might influence CTC spread from the bone tissue. Importantly, this correlation is not seen regarding treatment with the anti-bone resorption drug bisphosphonate, possibly because of different administration routes or dosing schedules [11] or, alternatively, potential off-target binding of denosumab to proteins other than receptor activator of nuclear factor κB ligand (RANKL).

Although its focus was on clinical parameters, our study did not provide molecular insights into the mechanism of action of denosumab in the context of its role in inhibiting CTC generation. Yet, considering that most denosumab-treated patients are characterized by bone metastatic disease but no primary breast tumor (which has been surgically removed prior to denosumab treatment), CTCs represent derivatives of their bone metastatic lesions. In this setting, we speculate that the effect of denosumab in suppressing CTC generation could be a result of RANKL inhibition within the bone, preventing the maturation of preosteoclasts into osteoclasts [12] and protecting the bone from degradation, leading to a lower likelihood of a bone metastatic lesion to shed CTCs. However, we cannot exclude an action of denosumab on breast cancer cells themselves, which have previously been shown to express high receptor activator of nuclear factor κB (RANK) levels [13, 14] and may be susceptible to its inhibition. Prospective studies and molecular assays will be needed to specifically dissect the role and mechanism of action of denosumab in CTC generation.

Recently, a phase 3 clinical trial designed to determine the long-term effects of denosumab treatment (D-CARE; ClinicalTrials.gov, NCT01077154) showed no benefits in metastasis-free survival and overall survival of patients with breast cancer. Importantly, individuals within this study were mainly patients with early breast cancer (i.e., stage IIB to IIIC), while our patient cohort was largely dominated by patients with stage IV disease. Although we are not aware of CTC enumeration data being evaluated within the D-CARE study, it is possible that denosumab might play a different role in the intravasation of bone metastasis-derived CTCs (as seen in our study) as opposed to primary tumor-derived CTCs (D-CARE).

Among other correlations, we observed an intriguing association between the absence of HER2 expression in the primary tumor and the presence of CTC clusters.

Although this result did not reach statistical significance, our observation regarding HER2 does not seem to be influenced by the metastatic tropism of HER2-positive breast cancers, and it might reveal important insights into the signaling networks involving CTC cluster formation, also considering HER2 expression fluctuations in CTCs and breast cancer metastasis [15, 16]. In other words, we speculate that HER2 signaling might influence cancer cells to intravasate as single CTCs, whereas its absence might point them toward collective invasion into the bloodstream. This hypothesis will require experimental testing.

We also found that CTC clusters, but not CTCs in general, are more prevalent in younger patients. Both CTC clusters and younger age have been associated with worse prognosis and reduced survival rates [2, 5, 17–20]. In this case, it is unlikely that younger age represents an independent risk factor for CTC cluster formation, but rather it may reflect an association with tumor aggressiveness [21].

Last, blood counts at the time of CTC collection provide evidence for applying well-established, cost-effective and widespread blood-testing strategies to stratify patients with higher likelihood to present with detectable CTCs. For instance, we find that lower RBC count has good correlation with the presence of CTCs. Additionally, CA 15-3 tumor antigen is highly increased in patients with CTC clusters, possibly reflecting a higher tumor load but also tumors that are characterized by an elevated shedding of mucin 1 (MUC-1)-containing cells into the bloodstream [22]. A functional relationship between MUC-1 and CTC clusters remains to be investigated. We also observed that higher MCV, higher MPV and higher WBC counts correlate with the presence of CTC clusters. We envision these parameters to be used to stratify patient populations to conduct CTC-related studies in the setting of advanced breast cancer.

Altogether, our study is meant as an exploratory analysis to evaluate the association of multiple clinical predictors with the presence of CTCs. Given the high number of hypotheses tested and the relatively low number of patient samples in the study ($n = 73$), none of the associations reported show a P value less than 0.05 after adjustment for multiple comparisons, with the exception of the correlation between denosumab treatment and the presence of bone metastasis (adjusted $P = 0.01$). For this reason, subsequent prospective and experimental studies should be conducted to validate the associations that are presented in this work, including the role of denosumab in CTC shedding.

Conclusions

Our data provide evidence of the association between treatment with the monoclonal antibody denosumab and the absence of CTCs from the peripheral circulation of patients with breast cancer. This finding suggests that denosumab treatment may be beneficial to reduce cancer spread in patients who are diagnosed with bone metastasis.

Although factors such as limited blood volume and diverse CTC isolation technologies may influence CTC detection rate in patients with cancer, the identification of a set of clinical correlatives to CTCs in breast cancer is likely to facilitate the identification of those patients who would benefit the most from CTC analysis, including genetic profile assessment for patient stratification [6, 23] and testing of drug susceptibility [7]. As an added benefit to this analysis, the identification of denosumab treatment as a strategy to reduce CTC intravasation warrants further investigation.

Additional files

Additional file 1: Table S1. Drug classification. The table shows the drug classification used for the analysis, grouping drugs into targeted therapy, chemotherapy, hormone therapy and immunotherapy. *Abbreviations: cbx6* Chromobox 6, *CDK4* Cyclin-dependent kinase 4, *CDK6* Cyclin-dependent kinase 6, *EGFR* Epidermal growth factor receptor, *Her2* Human epidermal growth factor receptor 2, *mTOR* Mechanistic target of rapamycin, *PD-L1* Programmed cell death 1 ligand 1, *VEGF* Vascular endothelial growth factor. (XLSX 10 kb)

Additional file 2: Table S2. Variable classification and statistical test applied. The table shows the type of variable and statistical test used for the analysis of individual clinicopathological parameters. *Abbreviations: ER*, Estrogen receptor, *HER2*, Human epidermal growth factor receptor 2, *PR* Progesterone receptor. (XLSX 10 kb)

Additional file 3: Table S3. Patient characteristics. The table shows the characteristics of the 73 patients included in the study. *Abbreviations: ER* Estrogen receptor, *HER2* Human epidermal growth factor receptor 2, *ID* Invasive ductal, *IL* Invasive lobular, *NA* Not available, *PR* Progesterone receptor, *ECOG* Eastern Cooperative Oncology Group (as defined by Oken et al. [24]). (XLSX 14 kb)

Additional file 4: Figure S1. Circulating tumor cell (CTC) capture strategy. (**a**) Schematic drawing showing the size-based capturing principle of the Parsortix microfluidic device (*left*). Plot showing the capture efficiency of the Parsortix microfluidic device for MCF7 cells spiked in healthy blood samples (*right*). Representative images of a captured MCF7 single cell, a cell cluster (*green*) and a contaminant white blood cell (WBC; red) in the Parsortix microfluidic cassette (*bottom*). (**b**) Representative images of a captured single CTC, a CTC cluster (*green*) and a contaminant WBC (*red*) from a breast cancer patient sample. (**c**) Bar graph showing the number of patients in whom no CTCs, single CTCs or CTC clusters were found. (PDF 245 kb)

Additional file 5: Table S4. Therapy evaluation in patients with circulating tumor cells. The table shows the types of therapy that patients with and without circulating tumor cells (CTCs) underwent. (XLSX 9 kb)

Additional file 6: Table S5. Bisphosphonates or denosumab treatment. The table shows whether bisphosphonates or denosumab was administered to each of the patients included in the study. (XLSX 10 kb)

Additional file 7: Table S6. Clinical features of patients who were treated or not with denosumab. The table shows clinical features of patients who were treated or not with denosumab. *Abbreviations: ER* Estrogen receptor, *HER2* Human epidermal growth factor receptor 2, *PR* Progesterone receptor. (XLSX 11 kb)

Additional file 8: Figure S2. Progression-free survival of patients who were treated or not with denosumab. Kaplan-Meier curve showing the progression-free survival probability of patients who were treated (*red*) or not (*green*) with denosumab (*top*). $P = 0.95$ by pairwise log-rank test. The table shows the number of patients at each time point (*bottom*). (PDF 188 kb)

Additional file 9: Table S7. Clinical features of patients with single circulating tumor cell and circulating tumor cell clusters. The table shows clinical features of patients in whom only single circulating tumor cells (CTC single cell) or also clustered circulating tumor cells (CTC clusters) were found. *Abbreviations: ER* Estrogen receptor, *HER2* Human epidermal growth factor receptor 2, *PR* Progesterone receptor. (XLSX 11 kb)

Additional file 10: Table S8. Complete blood counts in patients with single circulating tumor cells and circulating tumor cell clusters. The table shows complete blood counts in patients in whom only single circulating tumor cells (single CTC) or also clustered circulating tumor cells (CTC clusters) were found. *Abbreviations: CA 15-3* Cancer antigen 15-3, *CRP* C-reactive protein, *HCT* Hematocrit, *HGB* Hemoglobin, *LDH* Lactate dehydrogenase, *LUC* Large unstained cells, *MCH* Mean corpuscular hemoglobin, *MCHC* Mean corpuscular hemoglobin concentration, *MCV* Mean corpuscular volume, *MPV* Mean platelet volume, *PLT* Platelets, *RBC* Red blood cells, *WBC* White blood cells. (XLSX 10 kb)

Abbreviations

AF488: Alexa Fluor 488; CA 15-3: Cancer antigen 15-3; CRP: C-reactive protein; CTC: Circulating tumor cell; EDTA: Ethylenediaminetetraacetic acid; EGFR: Epidermal growth factor receptor; EpCAM: Epithelial cell adhesion molecule; ER: Estrogen receptor; FITC: Fluorescein isothiocyanate; HCT: Hematocrit; HER2: Human epidermal growth factor receptor 2; HGB: Hemoglobin; LDH: Lactate dehydrogenase; LUC: Large unstained cells; MCH: Mean corpuscular hemoglobin; MCHC: Mean corpuscular hemoglobin concentration; MCV: Mean corpuscular volume; MPV: Mean platelet volume; MUC-1: Mucin 1; PLT: Platelets; PR: Progesterone receptor; RANK: Receptor activator of nuclear factor κB; RANKL: Receptor activator of nuclear factor κB ligand; RBC: Red blood cells; WBC: White blood cells

Acknowledgements

We thank all the patients who participated in the study, all involved clinicians and study nurses, and all members of the Aceto laboratory for feedback and discussions. We also thank Julia Gutzwiller for data collection.

Funding

Research in the Aceto laboratory is supported by the European Research Council, the Swiss National Science Foundation, the Swiss Cancer League, the Basel Cancer League, the two cantons of Basel through the ETH Zürich, and the University of Basel.

Availability of data and materials

All data generated or analyzed during this study are included in this published article and its supplementary information files.

Authors' contributions

MV, JL, CB, AM, AZ, CK, VHS, WPW and CR consented all study patients, collected patient samples, interpreted the data and provided clinical input. BMS, SG, CD, IK and RS performed microfluidic CTC isolation and enumeration, data interpretation and troubleshooting. FCG. performed statistical analysis. NA supervised the study and drafted the manuscript. All authors read and approved the final version of the manuscript.

Ethics approval and consent to participate

All study participants agreed and signed a written informed consent. The study was performed under the protocols EKNZ BASEC 2016-00067 and EK 321/10, which received ethical and institutional review board approvals before study initiation (Ethics Committee northwest/central Switzerland [EKNZ]). This study was performed in compliance with the Declaration of Helsinki.

Consent for publication

Not applicable.

Competing interests

NA, BMS and FCG are listed as inventors in a patent application filed by the University of Basel, entitled "Inhibitors of bone resorption for treatment of metastasis" (#EP18161098). The other authors declare that they have no competing interests.

Publisher's Note

Author details

[1]Gynecologic Cancer Center, University Hospital Basel, 4056 Basel, Switzerland. [2]Department of Medical Oncology, University Hospital Basel, 4056 Basel, Switzerland. [3]Department of Biomedicine, Cancer Metastasis Laboratory, University of Basel and University Hospital Basel, Mattenstrasse 28, CH-4058 Basel, Switzerland. [4]SIB Swiss Institute of Bioinformatics, 1015 Lausanne, Switzerland. [5]Breast Center, University Hospital Basel, 4056 Basel, Switzerland.

Received: 5 April 2018 Accepted: 23 October 2018
Published online: 20 November 2018

References

1. Aceto N, Toner M, Maheswaran S, Haber DA. En route to metastasis: circulating tumor cell clusters and epithelial-to-mesenchymal transition. Trends Cancer. 2015;1:44–52.
2. Rack B, Schindlbeck C, Jückstock J, Andergassen U, Hepp P, Zwingers T, et al. Circulating tumor cells predict survival in early average-to-high risk breast cancer patients. J Natl Cancer Inst. 2014;106:dju066.
3. Cristofanilli M, Budd GT, Ellis MJ, Stopeck A, Matera J, Miller MC, et al. Circulating tumor cells, disease progression, and survival in metastatic breast cancer. N Engl J Med. 2004;351:781–91.
4. Giuliano M, Giordano A, Jackson S, Hess KR, De Giorgi U, Mego M, et al. Circulating tumor cells as prognostic and predictive markers in metastatic breast cancer patients receiving first-line systemic treatment. Breast Cancer Res. 2011;13:R67.
5. Aceto N, Bardia A, Miyamoto DT, Donaldson MC, Wittner BS, Spencer JA, et al. Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis. Cell. 2014;158:1110–22.
6. Carter L, Rothwell DG, Mesquita B, Smowton C, Leong HS, Fernandez-Gutierrez F, et al. Molecular analysis of circulating tumor cells identifies distinct copy-number profiles in patients with chemosensitive and chemorefractory small-cell lung cancer. Nat Med. 2017;23:114–9.
7. Yu M, Bardia A, Aceto N, Bersani F, Madden MW, Donaldson MC, et al. Cancer therapy: ex vivo culture of circulating breast tumor cells for individualized testing of drug susceptibility. Science. 2014;345:216–20.
8. Kaifi JT, Kunkel M, Dicker DT, Joude J, Allen JE, Das A, et al. Circulating tumor cell levels are elevated in colorectal cancer patients with high tumor burden in the liver. Cancer Biol Ther. 2015;16:690–8.
9. Cristofanilli M, Broglio KR, Guarneri V, Jackson S, Fritsche HA, Islam R, et al. Circulating tumor cells in metastatic breast cancer: biologic staging beyond tumor burden. Clin Breast Cancer. 2007;7:471–9.
10. Xu L, Mao X, Imrali A, Syed F, Mutsvangwa K, Berney D, et al. Optimization and evaluation of a novel size based circulating tumor cell isolation system. PLoS One. 2015;10:e0138032.
11. Whitaker M, Guo J, Kehoe T, Benson G. Bisphosphonates for osteoporosis—where do we go from here? N Engl J Med. 2012;366:2048–51.
12. Hanley DA, Adachi JD, Bell A, Brown V. Denosumab: mechanism of action and clinical outcomes. Int J Clin Pract. 2012;66:1139–46.
13. Blake ML, Tometsko M, Miller R, Jones JC, Dougall WC. RANK expression on breast cancer cells promotes skeletal metastasis. Clin Exp Metastasis. 2014; 31:233–45.
14. Pfitzner BM, Branstetter D, Loibl S, Denkert C, Lederer B, Schmitt WD, et al. RANK expression as a prognostic and predictive marker in breast cancer. Breast Cancer Res Treat. 2014;145:307–15.
15. Jordan NV, Bardia A, Wittner BS, Benes C, Ligorio M, Zheng Y, et al. HER2 expression identifies dynamic functional states within circulating breast cancer cells. Nature. 2016;537:102–6.
16. Houssami N, Macaskill P, Balleine RL, Bilous M, Pegram MD. HER2 discordance between primary breast cancer and its paired metastasis: tumor biology or test artefact? Insights through meta-analysis. Breast Cancer Res Treat. 2011;129:659–74.
17. Anderson WF, Pfeiffer RM, Dores GM, Sherman ME. Comparison of age distribution patterns for different histopathologic types of breast carcinoma. Cancer Epidemiol Biomark Prev. 2006;15:1899–905.
18. Adami HO, Malker B, Holmberg L, Persson I, Stone B. The relation between survival and age at diagnosis in breast cancer. N Engl J Med. 1986;315:559–63.

19. Fredholm H, Magnusson K, Lindstrom LS, Garmo H, Falt SE, Lindman H, et al. Long-term outcome in young women with breast cancer: a population-based study. Breast Cancer Res Treat. 2016;160:131–43.
20. Bleyer A, Barr R, Hayes-Lattin B, Thomas D, Ellis C, Anderson B, et al. The distinctive biology of cancer in adolescents and young adults. Nat Rev Cancer. 2008;8:288–98.
21. Koleckova M, Kolar Z, Ehrmann J, Korinkova G, Trojanec R. Age-associated prognostic and predictive biomarkers in patients with breast cancer. Oncol Lett. 2017;13:4201–7.
22. Duffy MJCA. 15-3 and related mucins as circulating markers in breast cancer. Ann Clin Biochem. 1999;36(Pt 5):579–86.
23. Lohr JG, Adalsteinsson VA, Cibulskis K, Choudhury AD, Rosenberg M, Cruz-Gordillo P, et al. Whole-exome sequencing of circulating tumor cells provides a window into metastatic prostate cancer. Nat Biotechnol. 2014;32:479–84.
24. Oken MM, Creech RH, Tormey DC, Horton J, Davis TE, McFadden ET, Carbone PP. Toxicity and response criteria of the Eastern Cooperative Oncology Group. Am J Clin Oncol. 1982;5(6):649–55.

Cell

Article

Circulating Tumor Cell Clustering Shapes DNA Methylation to Enable Metastasis Seeding

Sofia Gkountela,[1] Francesc Castro-Giner,[1,2] Barbara Maria Szczerba,[1] Marcus Vetter,[3,4] Julia Landin,[4,5] Ramona Scherrer,[1] Ilona Krol,[1] Manuel C. Scheidmann,[1] Christian Beisel,[6] Christian U. Stirnimann,[7] Christian Kurzeder,[3,5] Viola Heinzelmann-Schwarz,[3] Christoph Rochlitz,[4] Walter Paul Weber,[5] and Nicola Aceto[1,8,*]

[1]Cancer Metastasis Laboratory, Department of Biomedicine, University of Basel and University Hospital Basel, 4058 Basel, Switzerland
[2]SIB Swiss Institute of Bioinformatics, 1015 Lausanne, Switzerland
[3]Gynecologic Cancer Center, University Hospital Basel, 4056 Basel, Switzerland
[4]Department of Medical Oncology, University Hospital Basel, 4056 Basel, Switzerland
[5]Breast Center, University Hospital Basel and University of Basel, 4031 Basel, Switzerland
[6]Department of Biosystems Science and Engineering, ETH Zurich, 4058 Basel, Switzerland
[7]NEXUS Personalized Health Technologies, ETH Zurich, 8092 Zurich, Switzerland
[8]Lead Contact
*Correspondence: nicola.aceto@unibas.ch
https://doi.org/10.1016/j.cell.2018.11.046

SUMMARY

The ability of circulating tumor cells (CTCs) to form clusters has been linked to increased metastatic potential. Yet biological features and vulnerabilities of CTC clusters remain largely unknown. Here, we profile the DNA methylation landscape of single CTCs and CTC clusters from breast cancer patients and mouse models on a genome-wide scale. We find that binding sites for stemness- and proliferation-associated transcription factors are specifically hypomethylated in CTC clusters, including binding sites for OCT4, NANOG, SOX2, and SIN3A, paralleling embryonic stem cell biology. Among 2,486 FDA-approved compounds, we identify Na^+/K^+ ATPase inhibitors that enable the dissociation of CTC clusters into single cells, leading to DNA methylation remodeling at critical sites and metastasis suppression. Thus, our results link CTC clustering to specific changes in DNA methylation that promote stemness and metastasis and point to cluster-targeting compounds to suppress the spread of cancer.

INTRODUCTION

Circulating tumor cells (CTCs) are those cells that depart from cancerous lesions and enter the bloodstream (Alix-Panabières and Pantel, 2013). Although extraordinarily rare compared with blood cells and forced to strive for survival in circulation, CTCs are considered to be precursors of metastasis in various cancer types, including breast cancer (Aceto et al., 2015; Alix-Panabières and Pantel, 2014). CTCs are found in the blood of cancer patients as single CTCs and CTC clusters (Fidler, 1973; Liotta et al., 1976), with the latter featuring a higher ability to seed metastasis (Aceto et al., 2014). However, it is unknown what drives their enhanced metastatic potential and what are the vulnerabilities of clustered CTCs.

Abnormal DNA methylation patterns, including both genome-wide hypomethylation and hypermethylation, have been associated with several human cancers (Ehrlich, 2002, 2009; Feinberg et al., 2006; Klutstein et al., 2016; Lee et al., 2006). Generally, these cancer-associated epigenetic modifications appear to affect distinct genomic areas, with hypomethylation favoring regulatory and repetitive elements versus hypermethylation, which is more frequent in CpG islands (Ehrlich, 2002). Both modifications have the ability to alter the expression of neighboring genes and contribute to the cancer phenotype (Ehrlich, 2009; Klutstein et al., 2016). For regulatory elements, loss of DNA methylation at transcription factor binding sites (TFBSs) can designate active transcription factor networks or networks primed for activation at later stages, e.g., during processes such as the derivation of induced pluripotent stem cells from differentiated cells (Lee et al., 2014) or cancer progression (Feinberg and Vogelstein, 1983). Although DNA methylation analysis of primary tumors is extensively investigated (Feinberg and Vogelstein, 1983; Klutstein et al., 2016), the forces that shape the DNA methylome during metastatic dissemination are largely uncharacterized.

Here, we combine microfluidic-based CTC capture from breast cancer patients and mouse models, single-cell resolution DNA methylation and RNA expression analysis, a drug screen with 2,486 FDA-approved compounds, and functional validation studies in mouse models to gain insights into the biology and vulnerabilities of CTC clusters. Our study provides a genome-wide DNA methylation landscape of single and clustered CTCs in breast cancer, highlighting fundamental differences that affect metastasis and enabling the identification of cluster-targeting compounds with immediate clinical applicability.

RESULTS

Identification of Differentially Methylated Regions in CTC Clusters and Single CTCs

We first sought to identify active transcription factor networks by means of accessible TFBSs in single and clustered human

breast CTCs, matched within individual liquid biopsies, through a genome-wide single-cell resolution DNA methylation analysis. To this end, blood samples were drawn from 43 patients with progressive breast cancer and processed with the Parsortix device (Xu et al., 2015), a microfluidic technology that allows a size-based, antigen-agnostic enrichment of CTCs from unprocessed blood samples, specifically adapted to achieve a capture rate of >97.2% for single CTCs and >99.3% for CTC clusters, and no artificial cluster formation during sample processing (Figures S1A and S1B). Upon capture, live CTCs were stained for cell surface expression of EpCAM, HER2, and EGFR, and counterstained with antibodies against CD45 to identify contaminant leukocytes (Figure S1C). Upon staining verification, we identified matched single and clustered CTCs in 19% of the analyzed samples (8/43 patients), and a total of 18 marker-positive single CTCs and 29 marker-positive CTC clusters from four patients were individually micromanipulated and deposited in lysis buffer for single-cell resolution whole-genome bisulfite sequencing (Table S1) (Farlik et al., 2015, 2016). In parallel, we isolated spontaneously generated GFP-labeled single CTCs and CTC clusters from three mouse xenograft models, including two human breast CTC-derived cell lines (BR16 and BRx50) and the human breast cancer cell line MDA-MB 231 (lung metastatic variant, referred to as LM2) (Minn et al., 2005; Yu et al., 2014). In this setting, we individually micromanipulated 71 single CTCs and 48 CTC clusters (Table S1) and also processed them for single-cell resolution whole-genome bisulfite sequencing (Farlik et al., 2015, 2016). Samples with a low coverage (< 1,000 unique CpGs) or a low bisulfite conversion efficiency (CG/CHG/CHH < 97%)—corresponding to 10.7% of patient-derived samples and 0.8% of xenograft-derived samples—were excluded from the analysis, resulting in a total of 89 single CTCs and 71 CTC clusters from patients and xenografts. On average, we achieved 3.68% CpG coverage for single CTCs and 5.86% CpG coverage for CTC clusters, in line with recent single-cell whole-genome bisulfite sequencing studies (Farlik et al., 2015, 2016) (Figures S1D and S1E; Table S2). As expected, principal component analysis (PCA) mainly segregated CTCs based on the patient of origin or the specific xenograft model (Figures S1F and S1G). Metagene plot of CpG methylation revealed comparable methylation levels between single CTCs and CTC clusters across CpG islands, gene bodies, upstream (promoters) and downstream regions, including a drop of CpG methylation around the transcriptional start site, as expected (Figures S1H and S1I). We then specifically investigated differentially methylated regions (DMRs) between single CTCs and CTC clusters, evaluating average methylation levels in overlapping 5-kb windows, as previously established for single-cell DMR analysis (Farlik et al., 2015, 2016). For patient-derived CTCs, with this approach we identified 3,347 DMRs with a $\geq$ 80% methylation difference between single CTCs and CTC clusters. Of these, 1,305 regions were hypomethylated in CTC clusters and 2,042 were hypomethylated in single CTCs (Figure 1A). We then looked at xenograft-derived CTCs, and to evaluate a comparable number of DMRs as found in patients, we assessed overlapping regions with a $\geq$ 70% methylation difference between single CTCs and CTC clusters. We found a total of 1,430 DMRs, of which 909 hypomethylated in CTC clusters and 521 hypomethylated in single CTCs (Figure 1B). We then analyzed DMRs from both patient- and xenograft-derived CTCs using i-*cis*Target (Herrmann et al., 2012). With this analysis, among hypomethylated regions that are specific to either single CTCs or CTC clusters, we found a significant enrichment for several TFBSs, many of which overlapped between patient- and xenograft-derived CTCs (Figures 1C and 1D), thus allowing us to define specific hypomethylated TFBSs that globally characterize either single CTCs or CTC clusters in both patients and xenografts. Integrated gene ontology (GO) and pathway analysis of global CTC cluster hypomethylated TFBSs revealed a remarkable enrichment for stemness-related transcription factors that coordinately regulate proliferation and pluripotency, including OCT4, NANOG, SOX2, and SIN3A, paralleling embryonic stem cell (ESCs) biology (Figure 1E) (Kim et al., 2008; McDonel et al., 2012; Niwa, 2007; Saunders et al., 2017; van den Berg et al., 2010). Differently, single CTCs featured hypomethylation of other TFBSs, including those that are occupied by MEF2C, JUN, MIXL1, and SHOX2, commonly enriched in various cancers (Hong et al., 2014; Jiao et al., 2010; Laszlo et al., 2015; Raymond et al., 2014), yet independent of a core pluripotency network (Figure 1F) (Kim et al., 2008, 2010). To gain insights into more subtle changes in DNA methylation occurring specifically within promoters, gene bodies, and super enhancer regions, we carried out hypergeometric-based gene set enrichment analysis of genomic features in xenograft-derived CTCs (displaying a higher homogeneity compared to patient-derived CTCs). Consistently, this analysis revealed hypermethylation and H3K27me3 repression of Polycomb-repressive complex 2 (PRC2) target gene promoters and gene bodies (including those for SUZ12 and EED) in CTC clusters (Figures S1J–S1L), as previously alluded to in cancer specimens with stem-like and proliferative features (Avissar-Whiting et al., 2011; Kron et al., 2013; Lauss et al., 2012; Reddington et al., 2014; Wolff et al., 2010) and mirroring ESCs biology (Lee et al., 2006).

Thus, CTC clusters are clearly distinguishable from single CTCs based on their DNA methylation status at DMRs, where they mainly feature hypomethylation of bindings sites for stemness- and proliferation-associated TFs, such as OCT4, NANOG, SOX2, and SIN3A (Figure S1M, shown for patient CTCs), accompanied by a subtler hypermethylation of PRC2 target genes in promoters and gene bodies. This indicates that phenotypic differences in circulation affect DNA methylation dynamics.

Cluster-Associated Hypomethylated Regions Correlate with Poor Prognosis in Patients with Breast Cancer

We then tested whether the regions that are globally hypomethylated in CTC clusters are also hypomethylated in primary breast cancer. When analyzing bisulfite-sequencing data from The Cancer Genome Atlas (TCGA), we found 673/2,214 (30.39%) overlapping probes in 789 breast cancer patients, with 198 and 197 patients displaying either low (quantile Q1) or high (quantile Q4) methylation levels, respectively, in addition to a high correlation with genome-wide methylation levels (Figures 2A, S2A, and S2B). Progression-free survival (PFS) analysis on this subset of patients showed that low methylation levels within the regions that are hypomethylated in CTC clusters significantly correlate with a poor prognosis ($p < 0.05$) (Figure 2B). In contrast,

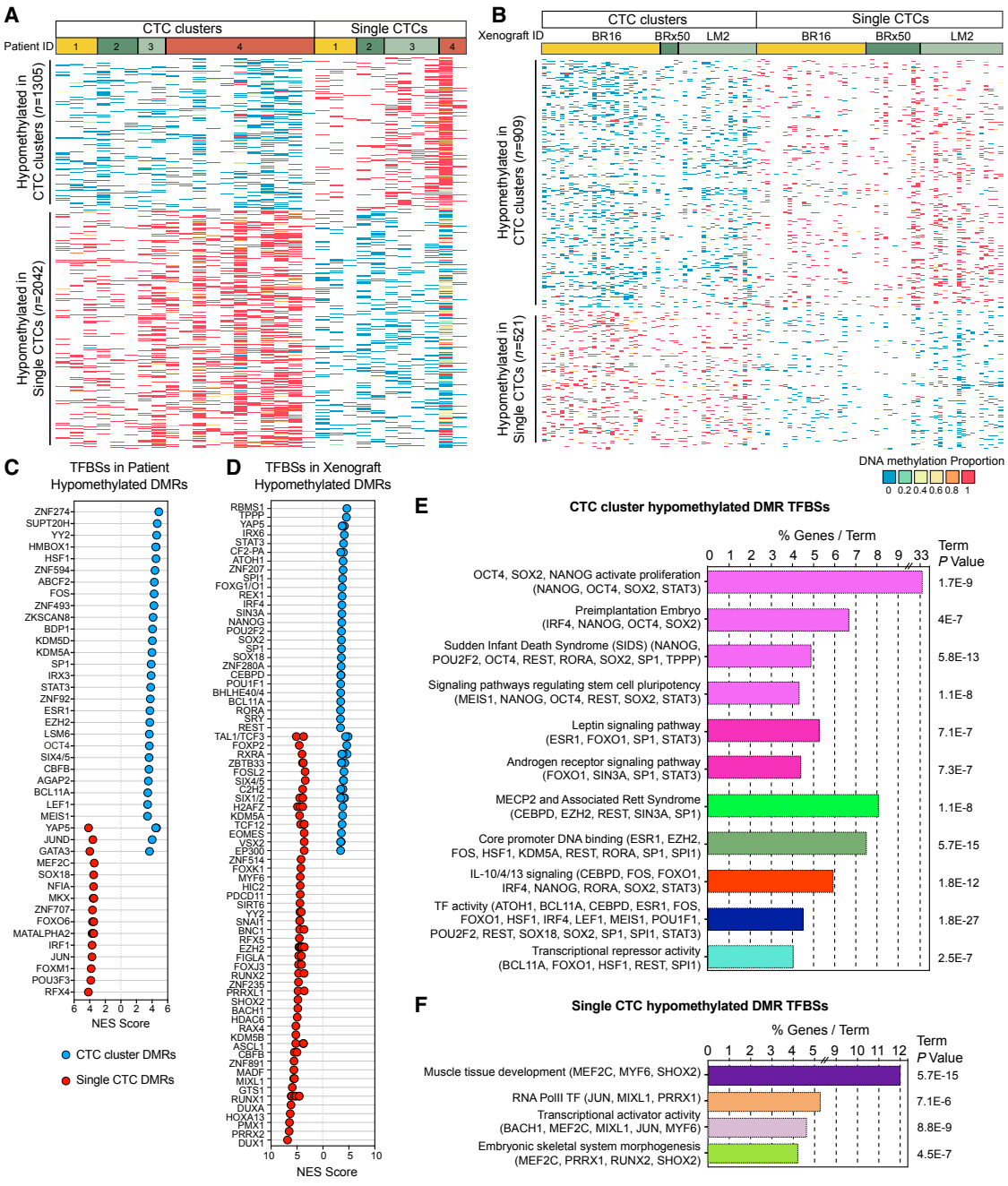


Figure 1. Whole-Genome Bisulfite Sequencing Analysis of CTCs from Breast Cancer Patients and Xenografts

(A) Heatmap showing methylation variable regions with ≥ 80% methylation difference between patient-derived CTC clusters and single CTCs (false discovery rate [FDR] < 0.05).

(B) Heatmap showing methylation variable regions with ≥ 70% methylation difference between xenograft-derived CTC clusters and single CTCs (FDR < 0.05).

(C and D) Normalized enrichment score (NES) representing enrichment (NES ≥ 3.4) of transcription factor binding sites (TFBSs) in CTC cluster hypomethylated regions (blue) and single CTC hypomethylated regions (red) of patients (C) or xenografts (D), identified using i-*cis*Target.

(E and F) Integrated gene ontology (GO) and pathway enrichment analysis of TFBSs identified using i-*cis*Target in hypomethylated regions of both patient- and xenograft-derived CTC clusters (E) or single CTCs (F). The bars represent the percentage of genes detected per GO and pathway term with p ≤ 0.05.

See also Figure S1 and Tables S1 and S2.

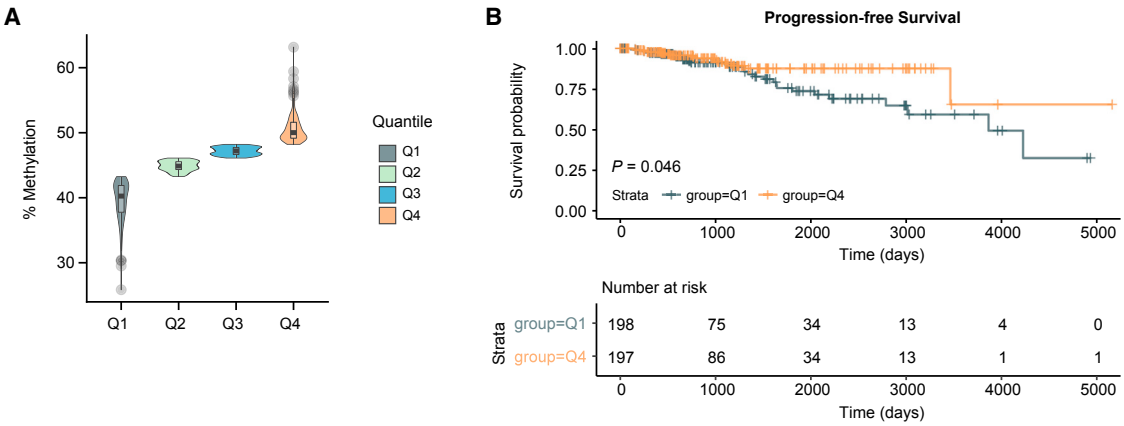


Figure 2. CTC Cluster Hypomethylated Regions Are Associated with a Poor Prognosis in Breast Cancer Patients

(A) Percentage of DNA methylation (mean beta values) of overlapping probes identified in the TCGA breast cancer patient dataset. Patients are grouped in four quantiles Q1-Q4, depending on the mean DNA methylation percentage of CTC cluster hypomethylated regions.

(B) Kaplan-Meier curve showing progression-free survival of breast cancer patients with overlapping probes in quantiles Q1 versus Q4 (top). The number of patients that progressed at each time point is shown (bottom).

See also Figure S2.

no correlation with PFS is observed for those regions that are hypomethylated in single CTCs (Figures S2C–S2F). Of note, while the majority of breast cancer patients in both quantile Q1 and Q4 belong to the hormone receptor-positive subtype (ER^+/PR^+), the proportion of patients with the HER2-positive subtype is significantly lower in quantile Q1 compared to quantile Q4 for both CTC cluster- and single CTC-associated hypomethylated regions (Figures S2G–S2J).

Together, our patient PFS analysis revealed that CTC cluster-associated hypomethylated regions are also hypomethylated at the level of the primary tumor in a subset of breast cancer patients characterized by a poor prognosis, compared to patients that display higher methylation levels in the same regions.

Proliferation-Related Genes Are Enriched in CTC Clusters

Our DNA methylation analysis led us to hypothesize that CTC clusters are characterized by active TF networks that support both stemness and proliferation. To identify whether the stemness- and proliferation-related TF networks are also transcriptionally active in CTC clusters compared to single CTCs, we performed single-cell resolution RNA sequencing analysis of 48 single CTCs and 24 CTC clusters, matched within individual liquid biopsies and isolated from six breast cancer patients with progressive metastatic disease, and of 49 single CTCs and 54 CTC clusters isolated from three xenograft models (Figures S3A and S3B; Tables S1 and S3). First, we performed a transcriptome-wide weighted gene co-expression network analysis (WGCNA) and identified 32 co-expression modules, revealing gene groups that are co-enriched in either single CTCs or CTC clusters (Figure 3A). Particularly, with this approach we identified two co-expression modules, containing total of 1,976 genes, enriched in patient-derived CTC clusters ($p < 0.02$; n = 1,544 for red; n = 432 for pink expression modules), while no co-expression modules were found to be significantly enriched in single CTCs (Figure 3B; Table S4). Gene ontology (GO) network analysis of the red and pink modules revealed an enrichment of gene groups related to cell-cell junctions, cellular proliferation, and platelet activation among others (Figure 3C). We then repeated the same analysis with xenograft-derived CTCs. Transcriptome-wide WGCNA identified 21 co-expression modules (Figure 3D), of which four significantly enriched in xenograft-derived CTC clusters and containing a total of 8,332 genes ($p < 0.003$; n = 159 for magenta, n = 337 for green, n = 753 for yellow, and n = 7,083 for turquoise expression modules), and three significantly enriched in xenograft-derived single CTCs and containing a total of 294 genes ($p < 0.005$; n = 202 for gray, n = 41 for dark turquoise, and n = 51 for dark red) (Figure 3E; Table S4). GO network analysis of the xenograft-derived CTC cluster networks revealed an enrichment of genes related to cell proliferation, DNA replication, cell-cell adhesion, and metabolic processes, among others (Figure 3F). In contrast, GO network analysis of the xenograft-derived single CTCs networks pointed to different processes that included RNA splicing, ATP metabolism, and ER stress response (Figure S3C). When specifically asking which processes were commonly found enriched in CTC clusters from both patients and xenografts, we identified genes related to both cellular proliferation and cell-cell adhesion (Figure 3G). While we previously reported upregulation of cell-cell adhesion components in CTC clusters (Aceto et al., 2014), whether cells within CTC clusters are also more proliferative compared to single CTCs is poorly understood. To directly address this point and validate our RNA sequencing findings, we stained CTCs from both patients and xenografts with the proliferation marker Ki67 and found that the percentage of Ki67-positive cells is indeed greatly increased in CTC clusters compared to matched single CTCs (Figures S3D–S3G).

In summary, our transcriptome-wide WGCNA revealed that CTC clusters—additionally to upregulating cell-cell junction components—are also characterized by a higher proliferation rate compared to single CTCs, in line with our DNA methylation results.

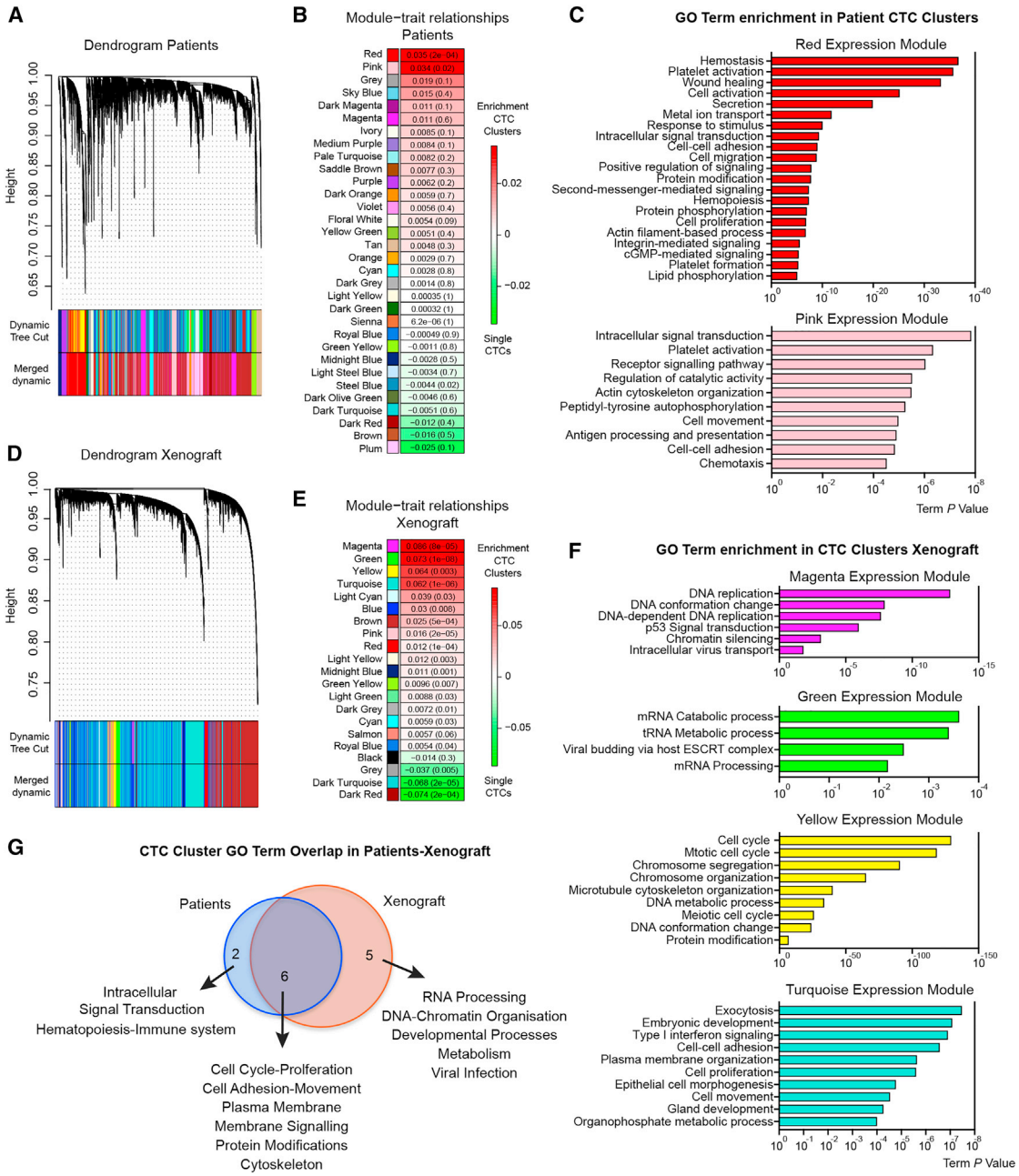


Figure 3. RNA Sequencing Analysis of Single CTCs and CTC Clusters from Breast Cancer Patients and Xenografts

(A) Weighted gene co-expression network analysis (WGCNA) in patient-derived single CTCs and CTC clusters, showing a hierarchical clustering tree of co-expression modules. Each module corresponds to a branch, which is labeled by a distinct color shown underneath.

(B) WGCNA identifies 32 modules with highly correlated gene expression patterns in patient-derived single CTCs and CTC clusters. Correlations between each module and CTC clusters or single CTCs are indicated by the intensity of red or green color, respectively. p value for each module is shown in brackets.

(C) GO term analysis of transcripts enriched in the red and pink expression modules of patient-derived CTC clusters.

(D) WGCNA in xenograft-derived single CTCs and CTC clusters, showing a hierarchical clustering tree of co-expression modules. Each module corresponds to a branch, labeled by a distinct color shown underneath.

(E) WGCNA identifies 21 modules with highly correlated gene expression patterns in xenograft-derived single CTCs and CTC clusters. Correlations between each module and CTC clusters or single CTCs are indicated by the intensity of red or green color, respectively. p value for each module is shown in brackets.

(F) GO term analysis of transcripts enriched in the magenta, green, yellow, and turquoise expression modules of xenograft-derived CTC clusters.

(G) Venn diagram showing the overlap of GO terms enriched in CTC clusters from patients and xenografts.

See also Figure S3 and Tables S3 and S4.

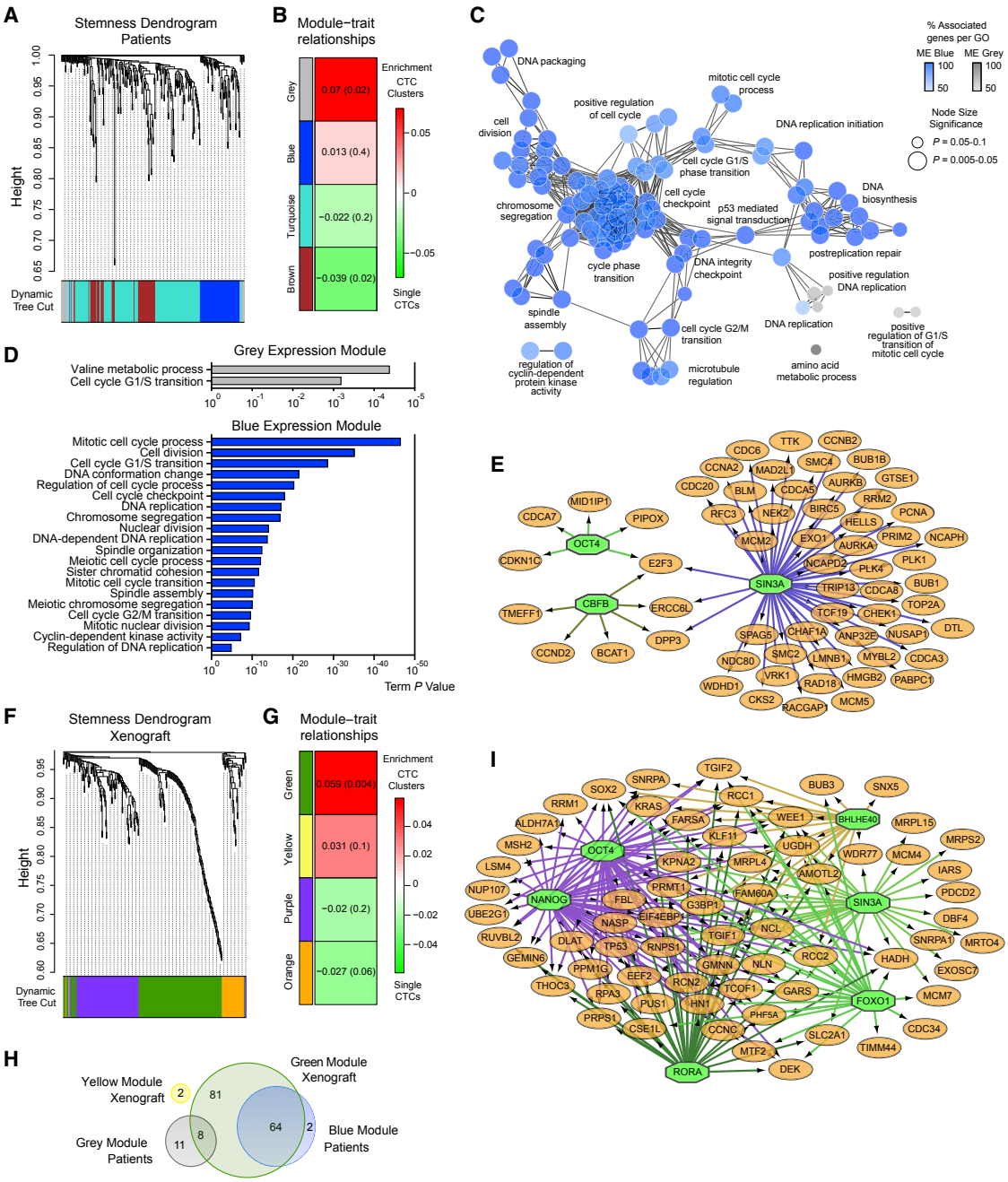


Figure 4. Stemness-Related Gene Expression Analysis of Single CTCs and CTC Clusters from Breast Cancer Patients and Xenografts

(A) WGCNA of 302 stemness-related transcripts in patient-derived single CTCs and CTC clusters, showing a hierarchical clustering tree of co-expression modules. Each module corresponds to a branch, which is labeled by a distinct color shown underneath.

(B) WGCNA identifies four modules with highly correlated gene expression patterns in patient-derived single CTCs and CTC clusters. Correlations between each module and CTC clusters or single CTCs are indicated by the intensity of red or green color, respectively. p value for each module is shown in brackets.

(C) GO network analysis of patient-derived transcripts identified in the CTC cluster-associated blue and gray modules using iRegulon. The node size indicates significance ($p < 0.05$), and color intensity corresponds to the percentage of genes that are associated to each GO category. Indicative GO categories are shown.

(D) GO term analysis of transcripts enriched in the blue and gray modules of patient-derived CTC clusters.

(E) Patient-derived CTC cluster-associated blue and gray module gene regulatory network analysis showing putative transcription factor dependence on SIN3A, OCT4, and CBFB (green octagons).

(F) WGCNA of 302 stemness-related transcripts in xenograft-derived single CTCs and CTC clusters, showing a hierarchical clustering tree of co-expression modules. Each module corresponds to a branch, labeled by a distinct color shown underneath.

(G) WGCNA identifies four modules with highly correlated gene expression patterns in xenograft-derived single CTCs and CTC clusters. Correlations between each module and CTC clusters or single CTCs are indicated by the intensity of red or green color, respectively. p value for each module is shown in brackets.

(legend continued on next page)

Stem-Cell-Related Genes Are Enriched in CTC Clusters

Our transcriptome-wide analysis revealed that the most significant differences between single and clustered CTCs involve the expression of cell-cell junction components and cell cycle progression. Yet our DNA methylation analysis pointed to both proliferation- and stemness-related TF networks being accessible in CTC clusters. To specifically ask whether CTC clusters would differ from single CTCs also in regard to the expression of stemness-related genes (which may be regulated with more subtle expression changes compared to cell cycle- and cell-cell junction-related genes), we focused our analysis on 335 genes that were previously shown to be consistently upregulated in mouse and human embryonic stem cells and embryonal carcinoma cells, as opposed to their differentiated counterparts (Wong et al., 2008), and asked whether they are co-enriched in CTC clusters. First, a subset of 302 of these 335 genes were found to be expressed in our CTC samples from breast cancer patients (cutoff $\geq$ 3 transcripts per million). WGCNA with these genes identified four expression modules co-enriched in either single CTCs or CTC clusters. Particularly, we identified two co-expression modules, containing a total of 85 genes, enriched in patient-derived CTC clusters (n = 66 for blue; n = 19 for gray expression module) and in two co-expression modules, containing 217 genes, enriched in patient-derived single CTCs (n = 156 for turquoise; n = 61 for brown) (Figures 4A and 4B; Table S4). GO network analysis of these modules confirmed that CTC clusters co-express stemness-related genes interconnected to cellular proliferation, while single CTCs co-express genes more related to metabolic processes (Figures 4C, 4D, and S4A; Table S4). Importantly, TF target gene analysis (Janky et al., 2014) of CTC cluster-enriched genes confirmed, among others, activity of SIN3A and OCT4, in line with our DNA methylation findings (Figure 4E; Table S5). We then extended our analysis of the 302 stemness-related genes to xenograft-derived CTCs. In this case, WGCNA revealed four expression modules, of which two were found to be enriched in xenograft-derived CTC clusters and containing a total of 153 genes (n = 151 for green; n = 2 for yellow expression module) (Figures 4F and 4G; Table S4), in high concordance (85% overlap) with patient-derived modules enriched in CTC clusters (Figure 4H; Table S4). TF target gene analysis of those genes that are enriched in xenograft CTC clusters also revealed the activity of several stemness- and proliferation-related TFs, including SIN3A, OCT4, NANOG, and SOX2 (Figure 4I; Table S5). Interestingly, the majority of these genes also displays hypomethylated promoter regions in CTC clusters compared to single CTCs (Figures S4B and S4C). Co-expression modules found in single CTCs from patients and xenografts also displayed significant overlap between each other (56.7%) (Figure S4D; Table S4), and TF analysis revealed the activity of MYF6 and ASCL1, also displaying hypomethylated binding sites in single CTCs (Figure S4E; Table S5).

Altogether, our gene expression data both at the transcriptome-wide level and also focused on stem cell-related genes strongly supports the model proposed with the DNA methylation analysis, suggesting that when compared to single CTCs, CTC clusters are endowed with a stemness- and proliferation-related network centered on the activity of key transcription factors including OCT4, SOX2, NANOG, and SIN3A. Activation of these programs may play a pivotal role in determining the metastasis-seeding ability of CTC clusters.

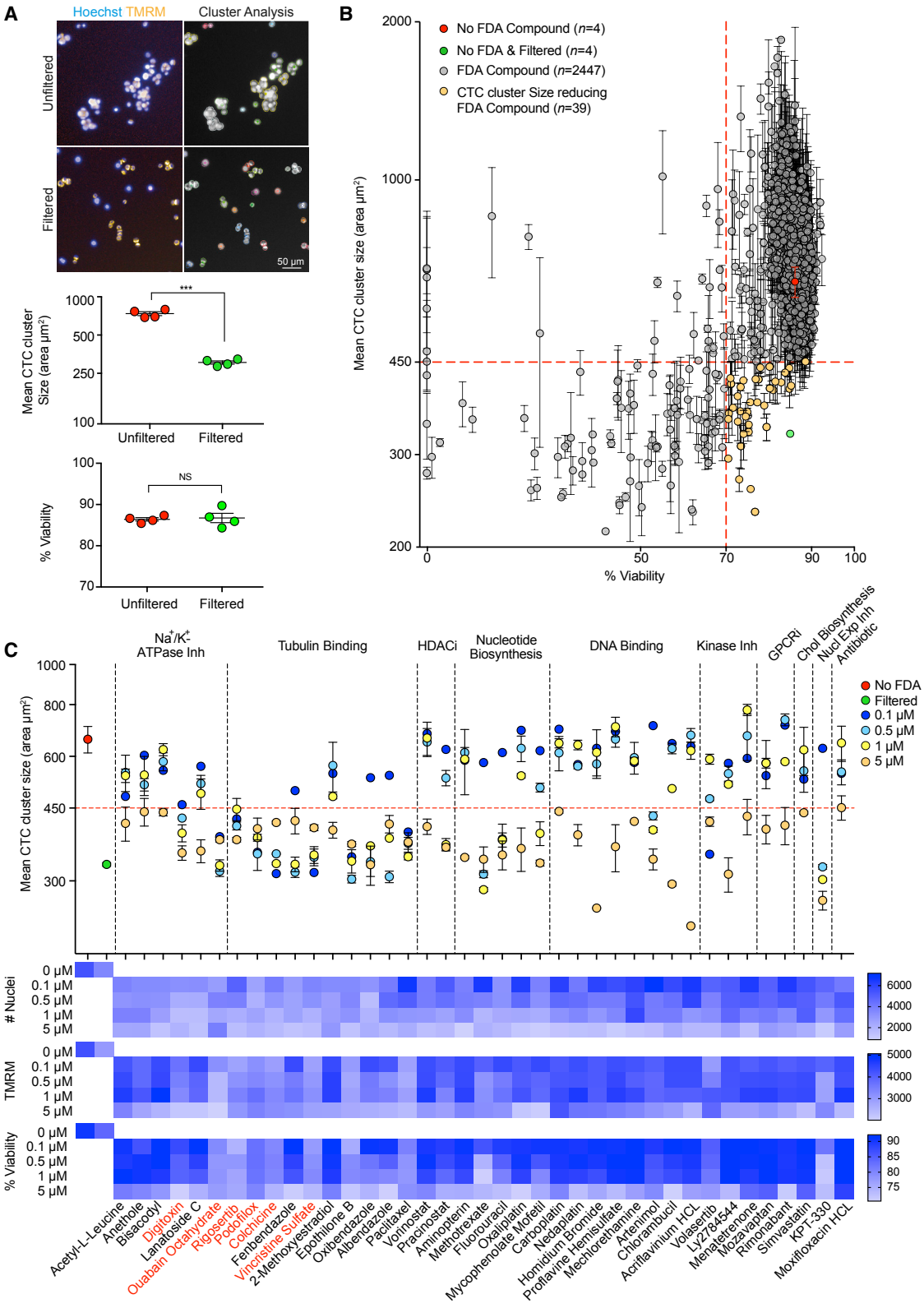
Identification of FDA-Approved Cluster-Targeting Agents

Next, we sought to identify actionable vulnerabilities of CTC clusters, and to test whether the epigenetic and transcriptional features of clustered CTCs are reversible upon cluster dissociation into single cells. To this end, we evaluated 2,486 FDA-approved compounds (Table S6) for their ability to dissociate clusters of human breast CTC-derived cells obtained from two patients with breast cancer (BR16 and BRx50), without affecting cellular viability. Cluster dissociation was assessed using an image-based high-content screening system and comparing CTCs treated with individual compounds to untreated CTCs as well as 40-μm-filtered single-cell suspension as negative and positive controls, respectively (Figures 5A, S5A, and S5B). For the majority of the 2,486 FDA-approved compounds, we observed no detectable reduction in cell viability (>70% viability) or in the mean CTC cluster size (> 450 μm^2) of CTC-derived cells upon treatment (gray circles; Figure 5B). Yet we identified 39 compounds that significantly ($p < 0.0001$) reduced mean CTC cluster size without compromising viability (orange circles; Figure 5B). These compounds include inhibitors of Na^+/K^+ ATPase (n = 6), histone deacetylase (HDAC) (n = 2), nucleotide biosynthesis (n = 5), kinases (n = 3), GPCRs (n = 2), cholesterol biosynthesis (n = 1), nuclear export (n = 1), tubulin (n = 10), DNA binding compounds (n = 8), and antibiotics (n = 1) (Figures 5B and S5C). Importantly, reducing compound concentration resulted in a gradual increase in mean CTC cluster size in both CTC-derived cell lines, followed by an increase in the number of nuclei detected, and a slight improvement in mitochondrial membrane potential and overall viability (Figures 5C and S5D). Under these conditions, six compounds consistently led to a significant decrease in mean CTC cluster size for both CTC-derived cell lines, even at the lowest concentrations tested. These compounds display similar functions and can be grouped into two families based on their mechanism of action, namely the Na^+/K^+-ATPase inhibitors digitoxin and ouabain octahydrate and the tubulin binding agents rigosertib, podofilox, colchicine, and vincristine sulfate (Figures 5C, S5C, and S5D).

Dissociation of CTC Clusters Leads to DNA Methylation Remodeling of Key Sites

To assess whether CTC cluster dissociation into single cells would lead to DNA methylation remodeling, we treated four CTC-derived cell lines with the six CTC cluster-dissociating compounds individually. We found that a prolonged treatment

(H) Venn diagram showing the overlap of CTC cluster-enriched transcripts between patients (blue and gray modules) and xenograft (yellow and green modules).
(I) Xenograft-derived CTC cluster-associated green and yellow module gene regulatory network analysis showing putative transcription factor dependence on SIN3A, OCT4, NANOG, BHLHE40, RORA, and FOXO1 (green octagons) and on SOX2 (orange circle).
See also Figure S4 and Tables S4 and S5.



(legend on next page)

(17 days) with 20 nM digitoxin or ouabain was unique in ensuring the highest degree of CTC cluster dissociation while allowing high viability and optimal proliferation rates comparable to untreated control cells (Figure S6A). Of note, the treatment period of 17 days was chosen to allow all cells (treated and untreated) to double at least 5 times and therefore enabling DNA methylation remodeling events to occur. Upon treatment, we processed BR16 and Brx50 cells for WGBS and RNA sequencing analysis to assess the molecular consequences of clusters dissociation. First, WGBS analysis revealed that the treatment did not affect global DNA methylation levels (data not shown). Yet a number of DMRs that were hypomethylated in CTC clusters gained methylation upon treatment (Figures S6B and S6C). Within the CTC cluster regions that gained methylation, i-*cis*Target analysis showed enrichment for binding sites of stemness-related TFs, such as OCT4, SOX2, and NANOG, and for SIN3A, among others (Figure 6A). Simultaneously, RNA expression analysis upon treatment with digitoxin or ouabain revealed a high concordance between the expression changes of stemness-related genes that were enriched in patient- and xenograft-derived CTC clusters ($r = 0.69$; $p < 2.22^{-16}$; Figures 6B), with the majority of these genes being downregulated as a consequence of the treatment ($p < 0.002$; Figure 6B). TF analysis revealed that downregulated genes were targets of OCT4, NANOG, SOX2, and SIN3A transcription factors, among others (Figure 6C). In contrast, TF analysis of those genes that were upregulated as a consequence to CTC cluster dissociation highlighted target genes of other TFs, independent from a core pluripotency network (Kim et al., 2008) (Figure S6D). Thus, CTC cluster dissociation into single cells with ouabain and digitoxin leads to DNA methylation remodeling and gain in methylation of critical binding sites for OCT4, SOX2, NANOG, and SIN3A, paralleled by downregulation of their corresponding target genes.

Inhibition of the Na^+/K^+ ATPase using ouabain and intracellular increase of calcium levels have been previously shown to negatively affect the formation of tight junctions and desmosomes in epithelial cells (Nigam et al., 1992; Rajasekaran et al., 2001; Stuart et al., 1994). Here, to gain more insights into the mechanism of action of Na^+/K^+ ATPase inhibitors in the context of CTC clusters disruption, we tested whether the suppression of ATP production and consequently, an increase in intracellular calcium levels—using the proton uncouplers FCCP and CCCP (Demaurex et al., 2009)—would also dissociate CTC clusters. Our results show that treatment of CTC-derived cells with increasing concentrations of FCCP and CCCP for 18 hr leads to a gradual decrease of mean cluster size (Figure 6D), paralleled by a significant increase in intracellular Ca^{2+} levels (Figure 6D). Interestingly, a similar gradual increase in intracellular Ca^{2+} is also observed when treating with increasing concentrations of digitoxin and ouabain (Figure S6E). These results support a model whereby intracellular Ca^{2+} increase leads to the inability of cancer cells to properly form cell-cell junctions.

To further confirm this model, we assessed whether cell-cell junction disruption in CTC-derived cells would lead to clusters dissociation as well as DNA methylation remodeling at CTC cluster-associated DMRs. To this end, we employed the CRISPR technology to simultaneously knockout both claudin 3 (CLDN3) and claudin 4 (CLDN4) in BR16 CTC-derived cells, two of the highest-expressed tight junction proteins in CTC clusters (Table S4). Using two independent sgRNAs for each gene, we generated three BR16 lines with double CLDN3/4 knockout, which also displayed a significant reduction of mean CTC cluster size (Figures S6F and S6G). Whole-genome bisulfite sequencing of the CLDN3/4 double-knockout cells showed that, upon dissociation into single cells and similarly to the events that occurred upon Na^+/K^+ ATPase inhibition, a number of CTC cluster-associated hypomethylated regions gained methylation (Figure 6E). Interestingly, i-*cis*Target analysis of the regions that gained higher levels of methylation revealed an enrichment of binding sites for OCT4, SOX2, NANOG, and SIN3A (Figure 6F), further indicating that CTC clustering directly affects DNA methylation dynamics at bindings sites for stemness- and proliferation-associated TFs.

Together, our results indicate that Na^+/K^+ ATPase inhibition leads to CTC clusters dissociation through the increase of the intracellular Ca^{2+} concentration and the consequent inhibition of cell-cell junction formation, resulting in DNA methylation remodeling at critical stemness- and proliferation-related binding sites.

Treatment with Na^+/K^+-ATPase Inhibitors Suppresses Spontaneous Metastasis Formation

To test whether ouabain and digitoxin would also enable CTC clusters disruption *in vivo*, we took a dual approach. First, we tested whether a 17-day *in vitro* treatment with ouabain and digitoxin would translate into a reduced ability of the treated cells to efficiently seed metastasis in untreated mice (Figure 7A). To this end, upon treatment, BR16 cells stably expressing GFP-luciferase were injected into the tail vein of NSG mice and

Figure 5. Screen for FDA-Approved Compounds that Dissociate CTC Clusters

(A) Representative images of unfiltered and filtered BR16 CTC-derived cells stained with Hoechst (blue) and TMRM (orange) (left). Representative images of single and clustered CTCs outline based on nuclei proximity as determined using the Colombus Image Analysis System (right). The plots show the mean CTC cluster size (area in micrometers squared) and percentage (%) of viability of unfiltered versus filtered BR16 cells ($n = 4$; ***$p < 0.001$ by Student's t test; ns, not significant) (bottom).

(B) Effect of a 2-day treatment of BR16 cells with 2,486 FDA-approved compounds at 5 μM concentration, plotted as mean CTC cluster size (area in micrometers squared) versus percentage (%) of viability ($n = 2$). Thirty-nine FDA-approved compounds (orange circles) result in significant decrease in mean CTC cluster size ($p < 0.0001$, F value = 7.71; DF = 38 using one-way ANOVA test; horizontal dashed red line, < 450 μm^2) and > 70% detectable viability (vertical-dashed red line). BR16 cells that were untreated (red) or 40 μm filtered (green) are shown as controls.

(C) The plot shows the mean CTC cluster size of BR16 cells treated with each of the 39 cluster-targeting compounds at four different concentrations. BR16 cells that were untreated (red) or 40 μm filtered (green) are shown as controls (top panel). The heatmap shows the number of nuclei, mean TMRM intensity, and percentage (%) of viability of BR16 cells treated with cluster-targeting compounds at the indicated concentrations (bottom). $n = 2$. Error bars represent SEM. See also Figure S5 and Table S6.

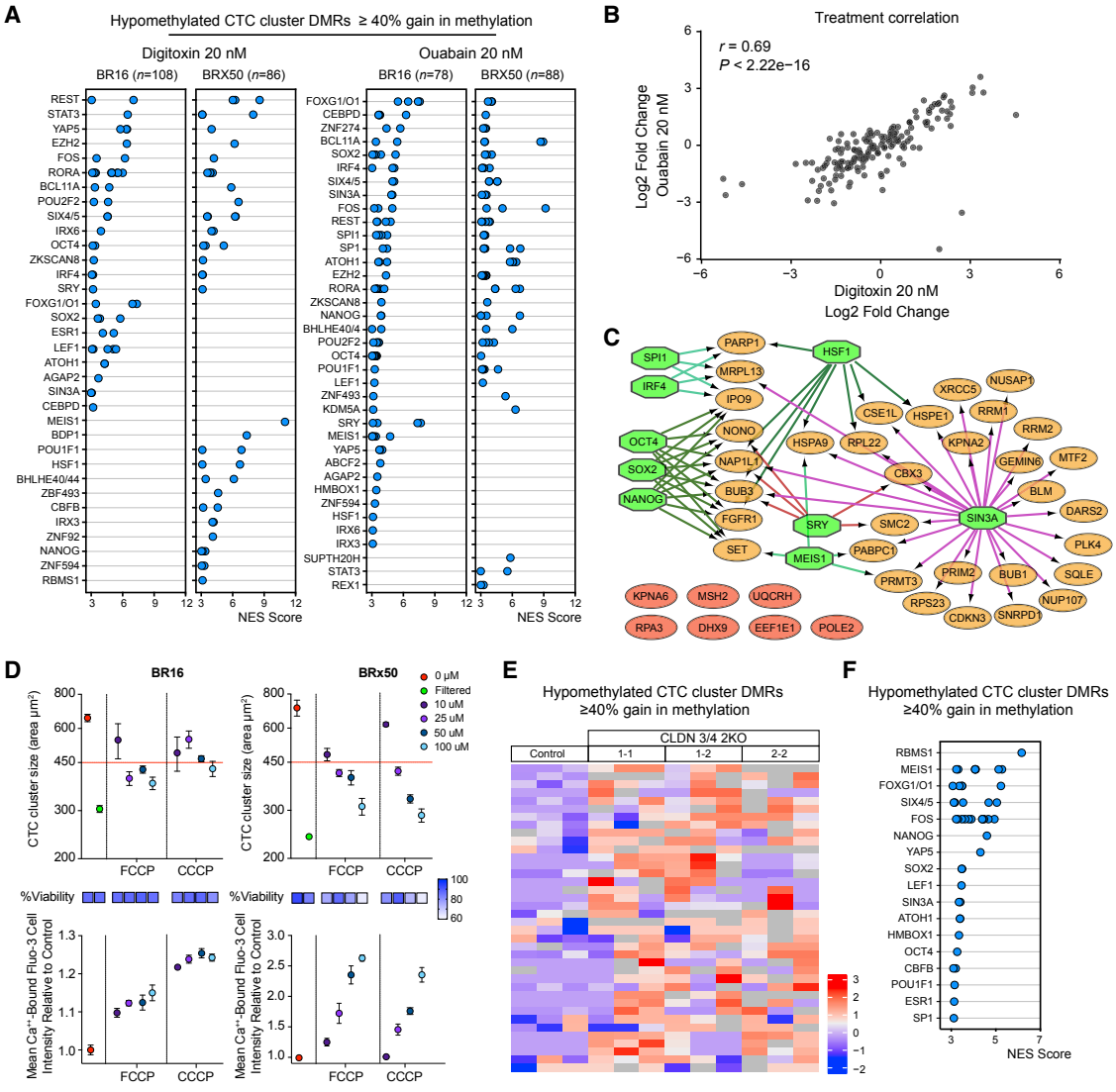


Figure 6. Treatment with Na^+/K^+ ATPase Inhibitors and Tight Junction Dissociation Restores Methylation at Key Sites

(A) i-*cis*Target analysis of CTC clusters hypomethylated regions, showing ≥ 40% methylation increase upon *in vitro* treatment with 20 nM of digitoxin or ouabain. Shown are enriched TFBSs in each CTC-derived cell line (NES ≥ 3).

(B) The plot shows expression changes in stemness-related transcripts that are enriched in patient- and xenograft-derived CTC clusters (n = 168) upon treatment of BR16 and BRx50 cells with 20 nM digitoxin or ouabain. Pearson's correlation coefficient (r) and p values are shown.

(C) Gene regulatory network analysis of downregulated genes upon treatment with 20 nM digitoxin or ouabain in both BR16 and BRx50 (log FC ≤ −0.5), showing putative dependence on OCT4, SOX2, NANOG, and SIN3A, among other TFs (green octagons). Downregulated genes that are not regulated by the TFs mentioned above are shown in red.

(D) The plot shows the mean CTC cluster size of BR16 and BRx50 CTC-derived cells treated with FCCP or CCCP at different concentrations (n = 4). BR16 and BRx50 cells that were untreated (red) or 40 µm filtered (green) are shown as controls (top). The heatmap shows the percentage (%) of viability of treated cells at the indicated concentrations (middle). The plot shows the mean cell intensity of Ca^{2+}-bound Fluo-3 after treatment with FCCP or CCCP, relative to the untreated control (red) (n = 4) (bottom). Error bars represent SEM.

(E) Heatmap showing methylation variable regions among CTC cluster hypomethylated regions that gain ≥ 40% methylation upon CLDN3/4 double knockout in BR16 cells.

(F) i-*cis*Target analysis of CTC cluster hypomethylated regions, showing ≥ 40% methylation increase upon CLDN3/4 double knockout in BR16 cells. Shown are enriched TFBSs (NES ≥ 3).

See also Figure S6 and Tables S2, S3, and S7.

noninvasively monitored through luminescence imaging for their ability to seed and propagate metastatic lesions. We found that while the treatment with digitoxin or ouabain did not affect the ability of BR16 cells to lodge in the lung tissue immediately after injection (see "day 0"; Figure 7B), it led to a reduced ability to survive during the first day upon arrival, as confirmed by a

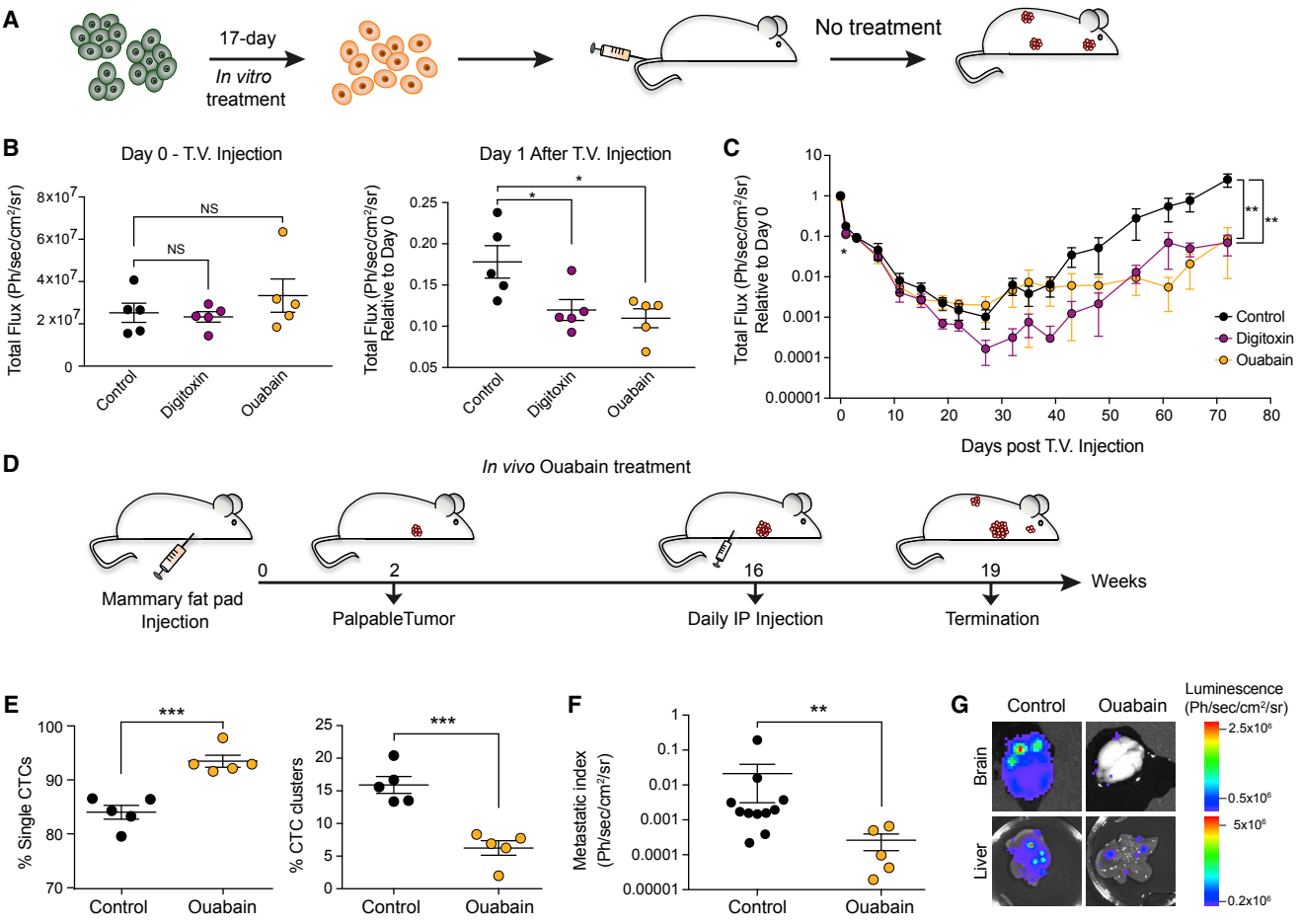


Figure 7. Treatment with Na^+/K^+ ATPase Inhibitors Suppresses Spontaneous Metastasis Formation

(A) Schematic representation of the experiment.

(B) The plots show the total bioluminescence flux at day 0 (left) and day 1 (right) upon tail vein injection of BR16 cells pre-treated with 20 nM digitoxin or ouabain. n = 5 for controls and ouabain, n = 4 for digitoxin; *p < 0.05 by Student's t test. ns, not significant. Error bars represent SEM.

(C) Metastasis growth curve over 72 days upon tail vein injection of BR16 cells pre-treated with 20 nM digitoxin or ouabain. n = 5; *p < 0.05; **p < 0.01 by Student's t test. Error bars represent SEM.

(D) Schematic representation of the experiment.

(E) The plots show the percentage (%) of spontaneously generated single CTCs and CTC clusters detected in the blood of BR16 xenografts treated with ouabain. n = 5; ***p < 0.001 by Student's t test. error bars represent SEM.

(F) The plot shows the metastatic index of BR16 xenografts treated with ouabain. n = 11 for controls, n = 5 for ouabain; **p < 0.01 by Student's t test. Error bars represent SEM.

(G) Representative images of the bioluminescence signal measured in brain and in liver of control and ouabain-treated NSG mice.

See also Figure S7.

significant increase in the expression of cleaved caspase 3 compared to control cells (see "day 1"; Figures 7B, S7A, and S7B). Overall, this difference in the ability to survive during the very early steps of metastasis seeding resulted in a delayed metastatic outgrowth despite the absence of further treatment *in vivo*, as measured during the course of 72 days upon injection (Figure 7C).

Second, mimicking more closely the clinical setting, to assess the effect of our CTC cluster-dissociation strategy for the spontaneous formation of CTC clusters and metastasis from a primary tumor, we injected BR16 cells in the mammary fat pad of NSG mice. Fourteen weeks after primary tumor formation we administered ouabain daily for three weeks and assessed CTC composition and the occurrence of spontaneous metastatic lesions (Figure 7D). Importantly, we observed that ouabain treatment reduced the frequency of spontaneously generated CTC clusters while increasing the frequency of single CTCs (Figure 7E), without altering the size of the primary tumor nor overall CTC numbers (Figures S7C and S7D). Along with a reduction in the frequency of CTC clusters, ouabain treatment also resulted in a remarkable suppression (80.7-fold) of the total metastatic burden (Figures 7F and 7G). In a similar fashion, when administering ouabain treatment to NSG mice carrying spontaneously metastasizing LM2 tumors, we also observed an increase in

the proportion of single CTCs and a decrease in CTC clusters (Figure S7E), without any change in the primary tumor size nor overall CTC numbers (Figures S7F and S7G), leading to a reduced metastatic burden compared to control (Figure S7H).

Together, these results demonstrate that Na^+/K^+ ATPase inhibition *in vivo* suppresses the ability of a cancerous lesion to spontaneously shed CTC clusters, leading to a remarkable reduction in metastasis seeding ability.

DISCUSSION

Our study provides a comprehensive genome-wide analysis of the DNA-methylation events that characterize CTCs in patients and xenografts. Surprisingly, we find that phenotypic differences—such as the ability of CTCs to navigate through the bloodstream as single cells or multicellular clusters—shape the DNA methylome. Clustering of CTCs results in hypomethylation of binding sites that are typically occupied by master stemness and proliferation regulators, including OCT4, NANOG, SOX2, and SIN3A, and hypermethylation of Polycomb target genes. More globally, we also find that the DNA methylation profile of CTC clusters is detected at the level of the primary tumor in a subset of breast cancers that are characterized by a poor prognosis. CTC clusters dissociation into single cells with Na^+/K^+ ATPase inhibitors or through cell-cell junction knockdown enables DNA methylation remodeling at critical sites, highlighting a direct connection between clustering and methylation status. As a result, Na^+/K^+ ATPase inhibitors treatment emerges as a new strategy to significantly reduce the spread of cancer, providing a rationale for using these compounds in clinical studies.

Our results suggest that CTC clusters may share several properties that commonly feature stem cell biology. For instance, OCT4, NANOG, SOX2, and SIN3A are predominantly active in embryonic stem cells (ESCs), simultaneously regulating self-renewal and proliferation (Kim et al., 2008; McDonel et al., 2012; Niwa, 2007; van den Berg et al., 2010). In addition, ESCs rely on Polycomb-mediated repression of differentiation genes and chromatin remodeling to maintain their active pluripotency network (Lee et al., 2006). Cell-cell junction activity has been shown in several instances to safeguard pluripotency and to be required for a complete reprogramming of somatic cells into stem cells, and disruption of cell-cell junctions (e.g., through targeting of E-cadherin) in human ESC results into OCT4, NANOG, and SOX2 downregulation along the loss of stemness features (Li et al., 2010a, 2010b, 2012; Pieters and van Roy, 2014). Thus, by analogy with stem cell biology, elevated expression of cell-cell junction components in cancer cells may not only enable their intravasation in the bloodstream as multicellular clusters but also their ability to retain stem-like features that facilitate metastasis initiation.

Identifying FDA-approved compounds that dissociate CTC clusters provides a new tool to reduce the spread of cancer. Our results demonstrate that CTC clusters disruption into single cells via the inhibition of the Na^+/K^+ ATPase has a dual, yet interconnected effect. The fact that clusters dissociation leads to DNA methylation remodeling at critical sites provides a direct link between a phenotypic state of the cells (i.e., clustered versus single) and DNA methylation dynamics. Second, CTC clusters disruption increases the proportion of single CTCs in the bloodstream but suppresses overall metastasis formation, indicating that targeting CTC clusters could be a valuable therapeutic strategy. In this regard, while we cannot exclude that an anti-cluster treatment might be beneficial also at the late disease stages in those patients whose metastasis may seed other metastases (McPherson et al., 2016; Reiter et al., 2017; Zhang et al., 2017), *in vivo* data suggest that treatment with ouabain and digitoxin should be administered early, ideally at the time of localized disease and before dissemination to distant sites, with the objective to prevent CTC cluster formation.

Ouabain and digitoxin are cardiac glycosides with a similar chemical structure, used in low doses for the treatment of hypotension and cardiac arrhythmias, acting through nonselective inhibition of the Na^+/K^+ ATPase (Altamirano et al., 2006; Schwartz, 1976). Upon inhibition of the Na^+/K^+ ATPase, cellular uptake of Na^+ occurs, leading to a simultaneous increase in intracellular Ca^{2+} and impaired translocation of desmosomal and tight junction proteins to the cellular membrane (Altamirano et al., 2006; Arispe et al., 2008). Based on this, our experimental evidences support a model whereby pharmacological inhibition of the Na^+/K^+ ATPase either directly or indirectly (i.e., by depletion of available ATP) leads to a concomitant increase in intracellular Ca^{2+} levels and CTC cluster disruption through suppression of functional cell-cell junction assembly (Cavey and Lecuit, 2009; Kim et al., 2011).

Together, our study provides key insights into the biology of CTCs and highlights a fundamental connection between phenotypic features of CTCs (such as their ability to circulate as multicellular clusters) and DNA methylation dynamics at critical stemness- and proliferation-related sites. Further, we identify the Na^+/K^+ ATPase inhibitors ouabain and digitoxin as FDA-approved agents capable to dissociate CTC clusters into single cells and to suppress spontaneous metastasis formation in xenograft models, providing a rationale for applying these compounds for the treatment of patients with breast cancer.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Human blood samples collection
  - Mouse blood samples collection
  - Cell lines
- METHOD DETAILS
  - Cell Culture
  - CTC Capture and Identification
  - Tumorigenesis Assays
  - Immunofluorescence Staining
  - FCCP, CCCP treatment and Intracellular Ca^{2+} measurement
  - CRISPR-CAS9 CLDN3/4 Double knock out in BR16

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures and seven tables and can be found with this article online at https://doi.org/10.1016/j.cell.2018.11.046.

ACKNOWLEDGMENTS

We thank all patients that participated in our study, all involved clinicians and study nurses, and all members of the Aceto lab for feedback and discussions. We thank Dr. Joan Massagué (MSKCC, New York); Dr. Daniel Haber and Dr. Shyamala Maheswaran (MGH and Harvard Medical School, Boston, MA) for donating cell lines; and Dr. Gerhard Christofori (University of Basel, Switzerland) for helpful comments on the manuscript. We thank Katja Eschbach and Elodie Burcklen from the Genomics Facility Basel (ETH Zürich) for sequencing support; Falco Kilchmann and Eva Maria Riegler from NEXUS PHT (ETH Zurich) for help with the drug screening; and Dr. Robert Ivanek (University of Basel) for support with data analysis. Calculations were performed at sciCORE scientific computing center at the University of Basel. Research in the Aceto lab is supported by the European Research Council (678834), the Swiss National Science Foundation (PP0P3_163938), the Swiss Cancer League (KFS-3811-02-2016 and KLS-4222-08-2017), the Basel Cancer League (KLbB-4173-03-2017), the two Cantons of Basel through the ETH Zürich (PMB-01-16), and the University of Basel.

AUTHOR CONTRIBUTIONS

S.G. and N.A. conceived the study, generated and interpreted the results, and wrote the manuscript. F.C.-G. conceived and performed the bioinformatics analyses. B.M.S. generated RNA sequencing (RNA-seq) libraries from CTCs. M.V., J.L., C.K. C.R., V.H.-S., and W.P.W. provided clinical input, consented breast cancer patients, and provided blood samples for CTC isolation. R.S. and I.K. performed cell culture experiments and contributed to the HTS compound screen. M.C.S. performed the CRISPR-related experiments. C.B. performed the WGBS and the RNA-seq library sequencing. C.U.S. provided the FDA-approved library for the drug screen. All authors approved the manuscript.

DECLARATION OF INTERESTS

N.A., S.G., and I.K. are listed as inventors in patent applications related to cancer treatment through CTC clusters inhibition. All other authors declare no competing interests.

Received: June 14, 2018
Revised: August 31, 2018
Accepted: November 28, 2018
Published: January 10, 2019

REFERENCES

Aceto, N., Bardia, A., Miyamoto, D.T., Donaldson, M.C., Wittner, B.S., Spencer, J.A., Yu, M., Pely, A., Engstrom, A., Zhu, H., et al. (2014). Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis. Cell *158*, 1110–1122.

Aceto, N., Toner, M., Maheswaran, S., and Haber, D.A. (2015). En route to metastasis: circulating tumor cell clusters and epithelial-to-mesenchymal transition. Trends Cancer *1*, 44–52.

Akalin, A., Kormaksson, M., Li, S., Garrett-Bakelman, F.E., Figueroa, M.E., Melnick, A., and Mason, C.E. (2012). methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. Genome Biol. *13*, R87.

Akalin, A., Franke, V., Vlahoviček, K., Mason, C.E., and Schübeler, D. (2015). Genomation: a toolkit to summarize, annotate and visualize genomic intervals. Bioinformatics *31*, 1127–1129.

Alix-Panabières, C., and Pantel, K. (2013). Circulating tumor cells: liquid biopsy of cancer. Clin. Chem. *59*, 110–118.

Alix-Panabières, C., and Pantel, K. (2014). Challenges in circulating tumour cell research. Nat. Rev. Cancer *14*, 623–631.

Altamirano, J., Li, Y., DeSantiago, J., Piacentino, V., 3rd, Houser, S.R., and Bers, D.M. (2006). The inotropic effect of cardioactive glycosides in ventricular myocytes requires Na+-Ca2+ exchanger function. J. Physiol. *575*, 845–854.

Arispe, N., Diaz, J.C., Simakova, O., and Pollard, H.B. (2008). Heart failure drug digitoxin induces calcium uptake into cells by forming transmembrane calcium channels. Proc. Natl. Acad. Sci. USA *105*, 2610–2615.

Avissar-Whiting, M., Koestler, D.C., Houseman, E.A., Christensen, B.C., Kelsey, K.T., and Marsit, C.J. (2011). Polycomb group genes are targets of aberrant DNA methylation in renal cell carcinoma. Epigenetics *6*, 703–709.

Cavey, M., and Lecuit, T. (2009). Molecular bases of cell-cell junctions stability and dynamics. Cold Spring Harb. Perspect. Biol. *1*, a002998.

Colaprico, A., Silva, T.C., Olsen, C., Garofano, L., Cava, C., Garolini, D., Sabedot, T.S., Malta, T.M., Pagnotta, S.M., Castiglioni, I., et al. (2016). TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. Nucleic Acids Res. *44*, e71.

Demaurex, N., Poburko, D., and Frieden, M. (2009). Regulation of plasma membrane calcium fluxes by mitochondria. Biochim. Biophys. Acta *1787*, 1383–1394.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21.

Edgar, R., Domrachev, M., and Lash, A.E. (2002). Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res. *30*, 207–210.

Ehrlich, M. (2002). DNA methylation in cancer: too much, but also too little. Oncogene *21*, 5400–5413.

Ehrlich, M. (2009). DNA hypomethylation in cancer cells. Epigenomics *1*, 239–259.

Farlik, M., Sheffield, N.C., Nuzzo, A., Datlinger, P., Schönegger, A., Klughammer, J., and Bock, C. (2015). Single-cell DNA methylome sequencing and bioinformatic inference of epigenomic cell-state dynamics. Cell Rep. *10*, 1386–1397.

Farlik, M., Halbritter, F., Müller, F., Choudry, F.A., Ebert, P., Klughammer, J., Farrow, S., Santoro, A., Ciaurro, V., Mathur, A., et al. (2016). DNA methylation dynamics of human hematopoietic stem cell differentiation. Cell Stem Cell *19*, 808–822.

Feinberg, A.P., and Vogelstein, B. (1983). Hypomethylation distinguishes genes of some human cancers from their normal counterparts. Nature *301*, 89–92.

Feinberg, A.P., Ohlsson, R., and Henikoff, S. (2006). The epigenetic progenitor origin of human cancer. Nat. Rev. Genet. *7*, 21–33.

Fidler, I.J. (1973). The relationship of embolic homogeneity, number, size and viability to the incidence of experimental metastasis. Eur. J. Cancer *9*, 223–227.

Herrmann, C., Van de Sande, B., Potier, D., and Aerts, S. (2012). i-cisTarget: an integrative genomics method for the prediction of regulatory features and cis-regulatory modules. Nucleic Acids Res. *40*, e114.

Hong, S., Noh, H., Teng, Y., Shao, J., Rehmani, H., Ding, H.F., Dong, Z., Su, S.B., Shi, H., Kim, J., et al. (2014). SHOX2 is a direct miR-375 target and a novel epithelial-to-mesenchymal transition inducer in breast cancer cells. Neoplasia *16*, 279–290.e1-5.

Huber, W., Carey, V.J., Gentleman, R., Anders, S., Carlson, M., Carvalho, B.S., Bravo, H.C., Davis, S., Gatto, L., Girke, T., et al. (2015). Orchestrating high-throughput genomic analysis with Bioconductor. Nat. Methods *12*, 115–121.

Janky, R., Verfaillie, A., Imrichová, H., Van de Sande, B., Standaert, L., Christiaens, V., Hulselmans, G., Herten, K., Naval Sanchez, M., Potier, D., et al. (2014). iRegulon: from a gene list to a gene regulatory network using large motif and track collections. PLoS Comput. Biol. *10*, e1003731.

Jiao, X., Katiyar, S., Willmarth, N.E., Liu, M., Ma, X., Flomenberg, N., Lisanti, M.P., and Pestell, R.G. (2010). c-Jun induces mammary epithelial cellular invasion and breast cancer stem cell expansion. J. Biol. Chem. *285*, 8218–8226.

Kim, J., Chu, J., Shen, X., Wang, J., and Orkin, S.H. (2008). An extended transcriptional network for pluripotency of embryonic stem cells. Cell *132*, 1049–1061.

Kim, J., Woo, A.J., Chu, J., Snow, J.W., Fujiwara, Y., Kim, C.G., Cantor, A.B., and Orkin, S.H. (2010). A Myc network accounts for similarities between embryonic stem and cancer cell transcription programs. Cell *143*, 313–324.

Kim, S.A., Tai, C.Y., Mok, L.P., Mosser, E.A., and Schuman, E.M. (2011). Calcium-dependent dynamics of cadherin interactions at cell-cell junctions. Proc. Natl. Acad. Sci. USA *108*, 9857–9862.

Klutstein, M., Nejman, D., Greenfield, R., and Cedar, H. (2016). DNA methylation in cancer and aging. Cancer Res. *76*, 3446–3450.

Kron, K., Trudel, D., Pethe, V., Briollais, L., Fleshner, N., van der Kwast, T., and Bapat, B. (2013). Altered DNA methylation landscapes of polycomb-repressed loci are associated with prostate cancer progression and ERG oncogene expression in prostate cancer. Clin. Cancer Res. *19*, 3450–3461.

Krueger, F., and Andrews, S.R. (2011). Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. Bioinformatics *27*, 1571–1572.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

Laszlo, G.S., Alonzo, T.A., Gudgeon, C.J., Harrington, K.H., Kentsis, A., Gerbing, R.B., Wang, Y.C., Ries, R.E., Raimondi, S.C., Hirsch, B.A., et al. (2015). High expression of myocyte enhancer factor 2C (MEF2C) is associated with adverse-risk features and poor outcome in pediatric acute myeloid leukemia: a report from the Children's Oncology Group. J. Hematol. Oncol. *8*, 115.

Lauss, M., Aine, M., Sjödahl, G., Veerla, S., Patschan, O., Gudjonsson, S., Chebil, G., Lövgren, K., Fernö, M., Månsson, W., et al. (2012). DNA methylation analyses of urothelial carcinoma reveal distinct epigenetic subtypes and an association between gene copy number and methylation status. Epigenetics *7*, 858–867.

Lee, T.I., Jenner, R.G., Boyer, L.A., Guenther, M.G., Levine, S.S., Kumar, R.M., Chevalier, B., Johnstone, S.E., Cole, M.F., Isono, K., et al. (2006). Control of developmental regulators by Polycomb in human embryonic stem cells. Cell *125*, 301–313.

Lee, D.S., Shin, J.Y., Tonge, P.D., Puri, M.C., Lee, S., Park, H., Lee, W.C., Hussein, S.M., Bleazard, T., Yun, J.Y., et al. (2014). An epigenomic roadmap to induced pluripotency reveals DNA methylation as a reprogramming modulator. Nat. Commun. *5*, 5619.

Li, D., Zhou, J., Wang, L., Shin, M.E., Su, P., Lei, X., Kuang, H., Guo, W., Yang, H., Cheng, L., et al. (2010a). Integrated biochemical and mechanical signals regulate multifaceted human embryonic stem cell functions. J. Cell Biol. *191*, 631–644.

Li, L., Wang, S., Jezierski, A., Moalim-Nour, L., Mohib, K., Parks, R.J., Retta, S.F., and Wang, L. (2010b). A unique interplay between Rap1 and E-cadherin in the endocytic pathway regulates self-renewal of human embryonic stem cells. Stem Cells *28*, 247–257.

Li, L., Bennett, S.A., and Wang, L. (2012). Role of E-cadherin and other cell adhesion molecules in survival and differentiation of human pluripotent stem cells. Cell Adhes. Migr. *6*, 59–70.

Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923–930.

Liotta, L.A., Saidel, M.G., and Kleinerman, J. (1976). The significance of hematogenous tumor cell clumps in the metastatic process. Cancer Res. *36*, 889–894.

Liu, J., Lichtenberg, T., Hoadley, K.A., Poisson, L.M., Lazar, A.J., Cherniack, A.D., Kovatich, A.J., Benz, C.C., Levine, D.A., Lee, A.V., et al. (2018). An integrated TCGA pan-cancer clinical data resource to drive high-quality survival outcome analytics. Cell *173*, 400–416.e11.

Lun, A.T., Bach, K., and Marioni, J.C. (2016). Pooling across cells to normalize single-cell RNA sequencing data with many zero counts. Genome Biol. *17*, 75.

Macaulay, I.C., Teng, M.J., Haerty, W., Kumar, P., Ponting, C.P., and Voet, T. (2016). Separation and parallel sequencing of the genomes and transcriptomes of single cells using G&T-seq. Nat. Protoc. *11*, 2081–2103.

Martin, H.L., Adams, M., Higgins, J., Bond, J., Morrison, E.E., Bell, S.M., Warriner, S., Nelson, A., and Tomlinson, D.C. (2014). High-content, high-throughput screening for the identification of cytotoxic compounds based on cell morphology and cell proliferation markers. PLoS ONE *9*, e88338.

McDonel, P., Demmers, J., Tan, D.W., Watt, F., and Hendrich, B.D. (2012). Sin3a is essential for the genome integrity and viability of pluripotent cells. Dev. Biol. *363*, 62–73.

McPherson, A., Roth, A., Laks, E., Masud, T., Bashashati, A., Zhang, A.W., Ha, G., Biele, J., Yap, D., Wan, A., et al. (2016). Divergent modes of clonal spread and intraperitoneal mixing in high-grade serous ovarian cancer. Nat. Genet. *48*, 758–767.

Minn, A.J., Gupta, G.P., Siegel, P.M., Bos, P.D., Shu, W., Giri, D.D., Viale, A., Olshen, A.B., Gerald, W.L., and Massagué, J. (2005). Genes that mediate breast cancer metastasis to lung. Nature *436*, 518–524.

Nigam, S.K., Rodriguez-Boulan, E., and Silver, R.B. (1992). Changes in intracellular calcium during the development of epithelial polarity and junctions. Proc. Natl. Acad. Sci. USA *89*, 6162–6166.

Niwa, H. (2007). How is pluripotency determined and maintained? Development *134*, 635–646.

Picelli, S., Faridani, O.R., Björklund, A.K., Winberg, G., Sagasser, S., and Sandberg, R. (2014). Full-length RNA-seq from single cells using Smart-seq2. Nat. Protoc. *9*, 171–181.

Pieters, T., and van Roy, F. (2014). Role of cell-cell adhesion complexes in embryonic stem cell biology. J. Cell Sci. *127*, 2603–2613.

Rajasekaran, S.A., Palmer, L.G., Moon, S.Y., Peralta Soler, A., Apodaca, G.L., Harper, J.F., Zheng, Y., and Rajasekaran, A.K. (2001). Na,K-ATPase activity is required for formation of tight junctions, desmosomes, and induction of polarity in epithelial cells. Mol. Biol. Cell *12*, 3717–3732.

Raymond, A., Liu, B., Liang, H., Wei, C., Guindani, M., Lu, Y., Liang, S., St John, L.S., Molldrem, J., and Nagarajan, L. (2014). A role for BMP-induced homeobox gene MIXL1 in acute myelogenous leukemia and identification of type I BMP receptor as a potential target for therapy. Oncotarget *5*, 12675–12693.

Reddington, J.P., Sproul, D., and Meehan, R.R. (2014). DNA methylation reprogramming in cancer: does it act by re-configuring the binding landscape of Polycomb repressive complexes? BioEssays *36*, 134–140.

Reiter, J.G., Makohon-Moore, A.P., Gerold, J.M., Bozic, I., Chatterjee, K., Iacobuzio-Donahue, C.A., Vogelstein, B., and Nowak, M.A. (2017). Reconstructing metastatic seeding patterns of human cancers. Nat. Commun. *8*, 14114.

Saunders, A., Huang, X., Fidalgo, M., Reimer, M.H., Jr., Faiola, F., Ding, J., Sánchez-Priego, C., Guallar, D., Sáenz, C., Li, D., and Wang, J. (2017). The

SIN3A/HDAC corepressor complex functionally cooperates with NANOG to promote pluripotency. Cell Rep. *18*, 1713–1726.

Schwartz, A. (1976). Is the cell membrane Na+, K+ -ATPase enzyme system the pharmacological receptor for digitalis? Circ. Res. *39*, 1–7.

Sheffield, N.C., and Bock, C. (2016). LOLA: enrichment analysis for genomic region sets and regulatory elements in R and Bioconductor. Bioinformatics *32*, 587–589.

Stuart, R.O., Sun, A., Panichas, M., Hebert, S.C., Brenner, B.M., and Nigam, S.K. (1994). Critical role for intracellular calcium in tight junction biogenesis. J. Cell. Physiol. *159*, 423–433.

van den Berg, D.L., Snoek, T., Mullin, N.P., Yates, A., Bezstarosti, K., Demmers, J., Chambers, I., and Poot, R.A. (2010). An Oct4-centered protein interaction network in embryonic stem cells. Cell Stem Cell *6*, 369–381.

Vocci, F.J., and London, E.D. (1997). Assessment of neurotoxicity from potential medications for drug abuse: ibogaine testing and brain imaging. Ann. N Y Acad. Sci. *820*, 29–39, discussion 39–40.

Wang, H.Q., Tuominen, L.K., and Tsai, C.J. (2011). SLIM: a sliding linear model for estimating the proportion of true null hypotheses in datasets with dependence structures. Bioinformatics *27*, 225–231.

Wang, L., Wang, S., and Li, W. (2012). RSeQC: quality control of RNA-seq experiments. Bioinformatics *28*, 2184–2185.

Wolff, E.M., Chihara, Y., Pan, F., Weisenberger, D.J., Siegmund, K.D., Sugano, K., Kawashima, K., Laird, P.W., Jones, P.A., and Liang, G. (2010). Unique DNA methylation patterns distinguish noninvasive and invasive urothelial cancers and establish an epigenetic field defect in premalignant tissue. Cancer Res. *70*, 8169–8178.

Wong, D.J., Liu, H., Ridky, T.W., Cassarino, D., Segal, E., and Chang, H.Y. (2008). Module map of stem cell genes guides creation of epithelial cancer stem cells. Cell Stem Cell *2*, 333–344.

Xu, L., Mao, X., Imrali, A., Syed, F., Mutsvangwa, K., Berney, D., Cathcart, P., Hines, J., Shamash, J., and Lu, Y.J. (2015). Optimization and evaluation of a novel size based circulating tumor cell isolation system. PLoS ONE *10*, e0138032.

Yu, M., Bardia, A., Aceto, N., Bersani, F., Madden, M.W., Donaldson, M.C., Desai, R., Zhu, H., Comaills, V., Zheng, Z., et al. (2014). Cancer therapy. Ex vivo culture of circulating breast tumor cells for individualized testing of drug susceptibility. Science *345*, 216–220.

Zhang, J., Cunningham, J.J., Brown, J.S., and Gatenby, R.A. (2017). Integrating evolutionary dynamics into treatment of metastatic castrate-resistant prostate cancer. Nat. Commun. *8*, 1816.

BJC British Journal of Cancer

www.nature.com/bjc

BRIEF COMMUNICATION

Genetics and Genomics

Detection of circulating tumour cell clusters in human glioblastoma

Ilona Krol[1], Francesc Castro-Giner[1,2], Martina Maurer[3], Sofia Gkountela[1], Barbara Maria Szczerba[1], Ramona Scherrer[1], Niamh Coleman[4], Suzanne Carreira[4], Felix Bachmann[3], Stephanie Anderson[3], Marc Engelhardt[3], Heidi Lane[3], Thomas Ronald Jeffry Evans[5], Ruth Plummer[6], Rebecca Kristeleit[7], Juanita Lopez[4] and Nicola Aceto[1]

Human glioblastoma (GBM) is a highly aggressive, invasive and hypervascularised malignant brain cancer. Individual circulating tumour cells (CTCs) are sporadically found in GBM patients, yet it is unclear whether multicellular CTC clusters are generated in this disease and whether they can bypass the physical hurdle of the blood–brain barrier. Here, we assessed CTC presence and composition at multiple time points in 13 patients with progressing GBM during an open-label phase 1/2a study with the microtubule inhibitor BAL101553. We observe CTC clusters ranging from 2 to 23 cells and present at multiple sampling time points in a GBM patient with pleomorphism and extensive necrosis, throughout disease progression. Exome sequencing of GBM CTC clusters highlights variants in 58 cancer-associated genes including *ATM*, *PMS2*, *POLE*, *APC*, *XPO1*, *TFRC*, *JAK2*, *ERBB4* and *ALK*. Together, our findings represent the first evidence of the presence of CTC clusters in GBM.

British Journal of Cancer (2018) 119:487–491; https://doi.org/10.1038/s41416-018-0186-7

INTRODUCTION

Human glioblastoma (GBM) is the most common and aggressive primary brain cancer in adults.[1] Yet, despite its characteristic invasive features and hypervascularity determined by angiogenic recruitment, only 0.4–2% of GBM patients develop metastasis outside of the central nervous system.[2] The rarity of these metastatic events has been attributed to the short-term survival of patients after initial GBM diagnosis, leaving insufficient time for establishment of extracranial lesions, as well as to the presence of the blood–brain barrier, which physically separates the brain from the rest of the body.[3] Circulating tumour cells (CTCs) are cancer cells that detach from a primary tumour lesion or a metastatic deposit and enter the bloodstream.[4] Although individual CTCs have been sporadically observed in GBM,[5–8] it is unclear whether multicellular CTC clusters[9] are generated and are able to pass through the blood-brain barrier in patients with GBM.

MATERIALS AND METHODS

Patient selection and recruitment

Patients were participants in the GBM arm of the ongoing study 'An open-label Phase 1/2a study of oral BAL101553 in adult patients with advanced solid tumours and in adult patients with recurrent or progressive glioblastoma or high-grade glioma' (NCT02490800). The main eligibility criteria for GBM patients in the study are measurable disease, defined by contrast-enhancing magnetic resonance imaging (MRI), and an Eastern Cooperative Oncology Group performance status of ≤ 2. The study was conducted in accordance with the Declaration of Helsinki (2000) and the International Conference on Harmonisation Guidelines for Good Clinical Practice. The study was approved by each study centre's Research Ethics Committee and all patients provided written informed consent before enrolment.

Mouse experiments

All mouse experiments were carried out according to institutional and cantonal guidelines (approved mouse protocol 2781, cantonal veterinary office of Basel-City). To measure GBM CTC intravascular aggregation, Nod Scid Gamma mice (The Jackson Laboratory, Bar Harbor, Maine, USA) were injected through the tail vein with a 1:1 mixture of T98G-GFP and T98G-RFP GBM cells, and upon injection, whole blood was withdrawn through a heart puncture and processed with the Parsortix device to characterise CTC composition.

GBM CTC enumeration

To test the GBM cell capture rate, healthy donor blood samples were drawn in three different collection tubes: CF DNA blood collection tube (BCT) (Streck, 218997), CellSave (CellSearch, 7900005) and Cyto-Chex BCT (Streck, 213386). The blood was then spiked with 300 human GBM T98G cells (Sigma, 92090213, CRL-1690) stably expressing green fluorescent protein (GFP).

[1]Department of Biomedicine, Cancer Metastasis Laboratory, University of Basel and University Hospital Basel, CH-4058 Basel, Switzerland; [2]Swiss Institute of Bioinformatics, Lausanne, Switzerland; [3]Basilea Pharmaceutica International Ltd., CH-4058 Basel, Switzerland; [4]Drug Development Unit, The Royal Marsden Hospital and The Institute of Cancer Research, London, UK; [5]Institute of Cancer Sciences, University of Glasgow, Glasgow, UK; [6]Northern Centre for Cancer Care, Newcastle upon Tyne, UK and [7]University College London Cancer Institute, London, UK

Correspondence: Nicola Aceto (Nicola.Aceto@unibas.ch)

These authors contributed equally: Ilona Krol, Francesc Castro-Giner.

Received: 2 February 2018 Revised: 12 June 2018 Accepted: 25 June 2018
Published online: 1 August 2018

Published by Springer Nature on behalf of Cancer Research UK

T98G-GFP cells were generated by transduction with UBC-GFP-T2A-Luc-expressing lentiviral particles (SBI, BLIV201PA-1-SBI). After incubation at room temperature (RT) for 24, 48, 72, and 96 h, samples were processed with the Parsortix microfluidic system for CTC enrichment. After wash with phosphate-buffered saline (PBS), Parsortix cassettes were scanned and GFP-positive cells were counted using Leica DMI 6000 microscope. To test artificial clustering of CTCs within the Parsortix cassettes, we spiked 100 single T98G-GFP and 100 single T98G-RFP into CF DNA BCT (Streck, 218997) tubes containing healthy donor blood. Upon processing through the Parsortix microfluidic system, we enumerated the ratios between single colour CTC and multicolour CTCs as a direct measure of on-chip CTC aggregation.

Tumour volume measurements

The enhancing tumour volume (cm^3) and the corresponding FLAIR signal abnormality volume (cm^3) in patients were calculated with Horos, an open source medical image viewer (LGPL license at Horosproject.org, sponsored by Nimble Co LLC d/b/a Purview in Annapolis, MD US).

CTC capture from GBM patients

Patients' blood samples were collected into the Streck CF DNA BCTs (Streck, 218997) and directly shipped at controlled temperature (RT). Samples were then loaded into the Parsortix microfluidic device within 24–48 h after collection from the patient and immediately processed for CTC enrichment. CTCs were then enriched from blood samples at 16 °C in disposable Parsortix cassettes (GEN3D6.5, ANGLE) according to the manufacturer's instruction and with a customised processing flow rate of 0.33 mm/s in the narrowest cassette gap (6.5 µm). After separation, GBM cells within cassettes were washed with PBS and further processed for immunostaining with antibodies against EGFR (Cell Signaling), Ki67 (Sigma), EB1 (kindly provided by Basilea Pharmaceutica), and CD45 (Thermo Fischer). As an additional confirmation, putative CTCs as well as a matched primary tumour biopsy of patient 4 were stained with antibodies against SOX2 (Cell Signaling).

Micromanipulation and exome sequencing

CTCs were released from the Parsortix cassette into ultra-low attachments culture plates. Cells were picked using micromanipulation (CellCelector, ALS) and transferred to RTL Plus buffer (Qiagen) within individual microcentrifuge tubes. Next, genomic DNA was amplified using the MDA method (GenomiPhi V3, GE Healthcare) and subject to Illumina library preparation with SureSelect XT Human All Exon V6 + Cosmic kit (Agilent Technologies). Sequencing was performed on HiSeq 2500 platform (Illumina) with 101 bp paired-end mode. For primary tumour sequencing, three 10 µm-thick paraffin sections from a primary tumour biospy were digested in DNA digestion buffer (50 mM Tris-HCl pH 8.5, 1 mM EDTA pH 8.0 and 0.5% Tween with addition of Proteinase K) in an eppendorf tube at 56 °C for 1 h, followed by 1 h incubation at 90 °C and 5 min at 95 °C. Samples were then kept on ice until RNase A treatment was performed (30 minutes at 37 °C). Next, DNA was precipitated using 7.5 M ammonium acetate and 100% isopropanol. After washing with 70% ethanol, the DNA pellet was air dried and then resuspended in standard Tris-containing buffer. Sequencing was performed on HiSeq 2500 platform (Illumina).

Exome-sequencing data analysis and mutation calling

After quality control performed with FastQC (v0.11.4), reads were mapped to the GRCh38 human reference genome using BWA-mem (v0.7.13) algorithm and sorted using Samtools (v1.3.1). Reads were then processed using Picard MarkDuplicates (v2.9.0) to remove duplicated reads and realigned using GATK IndelRealignment (v3.7.0) to improve alignment accuracy around indels. We used the interactive platform Ginkgo (http://qb.cshl.edu/ginkgo) to compute and plot the copy number profiles of single-cells. Input bed files were generated from BAM alignments using BEDTools (v2.26.0) and genomic coordinates were converted from GRCh38 to GRCh37 UCSC liftOver tool. Ginkgo was run with a using variable bin size of 500 kb simulated with 101 bp with BWA and normalised read counts for cell segmentation. Variants were identified using Monovar (20160514 update) for low-input DNA samples and bcftools (version 1.6) for primary GBM of patient 4. Resulting variants were annotated using snpEff (v4.3p) with Ensembl GRCh38.86 gene models. Putative somatic SNV were defined as variants not present in the white blood cell (WBC) pool and whose frequency in the dbSNP (build 150) was < 1%. In addition, a more stringent set of somatic single nucleotide variants (SNVs) was obtained by filtering out loci that were covered with < 5 reads in the WBC pool. Candidate somatic driver variants were defined as variants annotated with high or moderate impact according to snpEff, previously reported in the COSMIC (version 81) or located in recurrently mutated genes in gliobastoma according to the CGC database (version 81). Data analysis was conducted in R (version 3.4.0). Data visualisation was performed with the R package UpSetR and the Bioconductor package GenVisR.

RESULTS

CTC detection strategy in GBM patients

Patients selected for CTCs investigation were part of an ongoing open-label Phase 1/2a study arm of oral BAL101553, a water-soluble microtubule inhibitor, in adult individuals with recurrent or progressive GBM or high-grade glioma. Upon informed consent, serial peripheral blood samples were drawn in Streck tubes from 13 GBM patients at time points related to BAL101553 administration, including (1) within the 2 weeks before the first study-drug treatment; (2) pre-dose, (3) 2 h post-dose and (4) 24 h post-dose on Cycle 1 Day 1; and pre-dose on (5) Cycle 1 Day 8, (6) Cycle 1 Day 22 and (7) Cycle 2 Day 22 (Fig. 1a). Blood samples were then processed within 48 h for CTC enumeration and characterisation. To investigate CTC number and composition in GBM patients, we made use of a commercially available CTC detection method, the Parsortix microfluidic technology, customised to detect both single and clustered GBM CTCs in an antigen-independent manner and at > 98% capture efficiency, up to 48 h upon fixation in Streck CF DNA tubes (see methods and Supplementary Fig. 1a-c), and without the occurrence of artificial CTC aggregates during processing (Supplementary Fig. 1d,e). Briefly, blood samples were run through the Parsortix microfluidic cassette at a flow rate corresponding to 0.33 ml/s in the narrowest channel section, ensuring label-free physical capture of CTCs and their separation from blood components (Fig. 1b). Upon capture, CTCs were stained with a dedicated GBM CTC antibody cocktail containing antibodies against EGFR, Ki67 and the microtubule-associated protein EB1, as well as against the WBC marker CD45 to exclude leukocyte contamination (Supplementary Fig. 2). Putative GBM CTCs were scored positive when corresponding to at least one of these criteria: (1) cell diameter of at least 9 µm and negative CD45 staining, and/or (2) positive EGFR, Ki67 or EB1 staining and negative CD45 staining. With these parameters, and using healthy blood donors ($n = 3$) to define a false-positive detection threshold, we found that 7/13 patients were positive for at least 3 putative CTCs per 10 ml of blood during at least 1 time point (Fig. 1c). Importantly, radiological imaging determined that none of the 13 patients enroled in the study developed extracranial metastasis (data not shown), and no association was observed between functional MRI volume and the presence of CTCs, most likely due to the small patient cohort ($n = 13$) (Supplementary Fig. 3).

Detection of GBM CTC clusters

Among patients that resulted positive for putative GBM CTCs, we focused on a 43-year-old female patient (4), featuring the highest

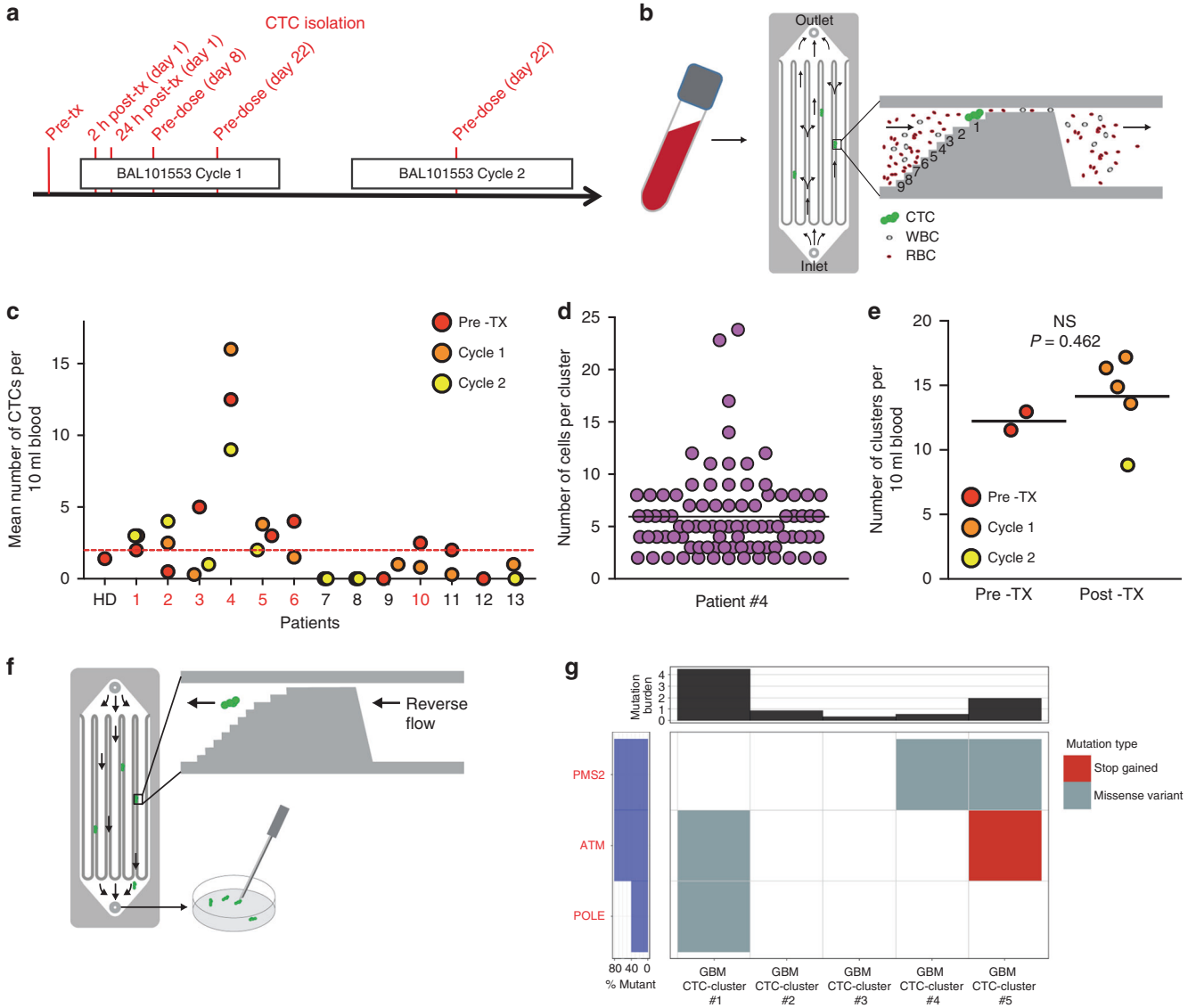


Fig. 1 Identification of GBM CTC clusters. **a** Schematic overview of the CTC isolation strategy (red) in GBM patients undergoing treatment with BAL101553. Each patient underwent several blood draws: one or two blood draws before BAL101553 treatment (pre-tx), four during the first BAL101553 cycle and one during the second BAL101553 cycle. **b** Schematic representation of our CTC enrichment strategy with the Parsortix microfluidic device. CTCs are captured with a size-based antigen-agnostic approach, while red and white blood cells flow through the device. **c** Plot showing the mean number of CTCs found at each sampling time point (pre-treatment: pre-TX; cycle 1 and cycle 2) in 13 GBM patients. Healthy donors samples (HD; $n = 3$) provide a false-positive threshold of two putative CTCs per 10 ml blood. Patients with > 2 putative CTCs per 10 ml of blood during at least one time point are shown in red. **d** Plot showing the number of cells per each GBM CTC cluster isolated from patient 4. **e** Plot showing the number of GBM CTC clusters per 10 ml of blood identified in patient 4 during each blood draw, both pre-treatment (pre-TX) and post-treatment (cycle 1 and cycle 2) with BAL101553. $P = 0.462$ by Student's *t*-Test. NS: not significant. **f** Schematic representation of the CTC isolation strategy by flow reversion and picking with a micromanipulator. **g** Putative somatic mutations in glioblastoma candidate driver genes found in GBM CTC clusters. The central panel shows the mutated genes coloured by the predicted consequence of the mutations. The top panel shows the number of mutations per sample and the left panel shows the percentage of samples that have a mutation in each represented gene

number of CTC events in the peripheral circulation. The patient had a first surgical resection of a right parietal GBM, which showed a high-grade astrocytic tumour with pleomorphism, mitoses, extensive necrosis and prominent vascular proliferation. Following surgery, the patient underwent chemoradiotherapy comprising standard radiation plus concomitant and subsequent adjuvant temozolomide. Imaging at the end of adjuvant treatment suggested progressive disease, requiring a second surgical resection followed by procarbazine/lomustine/vincristine (PCV) chemotherapy and re-irradiation. Further disease progression was noted and subsequently the patient was enroled onto the BAL101553 clinical trial (Supplementary Fig. 4a). MRI of the brain before commencing BAL101553 therapy and CTC enumeration showed enhancing tumour areas in the posterior aspect of the right parietal lobe, infiltrating the corpus callosum and extending periventricular along the occipital horn of the right lateral ventricle and inferiorly into the temporal lobe (Supplementary Fig. 4b). There were no other abnormalities seen on baseline whole-body radiological imaging (data not shown). CTC analysis revealed the presence of clusters of putative CTCs in the blood of this patient, ranging from 2 to 23 cells and expressing various combinations of EGFR, Ki67 and EB1 markers, yet negative for CD45 (Fig. 1d and Supplementary Fig. 4c). As a further validation step, putative GBM CTC clusters, as well as a biopsy from the primary GBM of patient 4

were stained with the GBM marker SOX2, revealing SOX2-positive cells in both specimens (Supplementary Fig. 4d, e). Similar to breast cancer,[9] we also confirmed that GBM CTCs are not able to form intravascular aggregates upon injection in the venous circulation of mice (Supplementary Fig. 5a,b), indicating that the putative CTC clusters observed in patient 4 are likely to be direct derivatives of the primary GBM tumour. As a relatively stable number of putative CTC clusters were observed in the samples taken prior to and after initiation of BAL101553 therapy, it does not appear as though the treatment itself impacted on CTC clusters release at this dose level (Fig. 1e).

Exome sequencing of GBM CTC clusters

To assess whether the cell clusters isolated from the patient of interest were bona fide GBM cells, we released them in solution from the Parsortix device and, with a semi-automatic micromanipulator, we deposited each cluster individually into lysis buffer (Fig. 1f), followed by DNA amplification and library preparation for exome sequencing. Six samples were successfully sequenced, including one WBC pool as germline control and five putative GBM CTC clusters. We obtained an average of 41 million reads per sample, corresponding to a median coverage depth of 32.6 × and an average breadth of 18.7%, in range with typical performance of single-cell-resolution exome sequencing of fixed cells[10] (Supplementary Fig. 6a and Suppl. Table 1). To address whether our putative GBM CTC clusters were indeed of cancer origin, we adopted two parallel approaches. First, we inferred DNA copy number variations (CNVs) from exome-sequencing data in all our samples (see Methods section). In general, while CNV assessment from single-cell-resolution exome sequencing cannot precisely determine CNV in exact regions due to the relative low coverage and possible amplification biases, we sought to use it to investigate general aneuploidy trends across cells. We found that while predicted ploidy for WBC control pool was 2, all putative GBM CTC clusters displayed a predicted ploidy ranging from 2.2 to 4.1, suggesting a higher degree of genomic rearrangements compared with the WBC control pool (Supplementary Fig. 6b). Second, we identified variations at the nucleotide level in both the WBC control pool as well as the putative GBM CTC clusters. Genetic variations identified in both the putative CTC clusters and the WBC control pool were considered likely to be part of the patient's germline variants and not included for further analyses. We then interrogated specific GBM-associated genes (see Methods section) and found among them a premature STOP codon in *ATM* gene and missense variants in *PMS2* and *POLE* genes in GBM CTC clusters (Fig. 1g and Suppl. Table 2). Of note, the region corresponding to the STOP codon in *ATM* gene and the majority of the other variants were covered and found to contain a wild-type sequence in the WBC pool control sample (Suppl. Table 2). More generally, when interrogating the whole exome of the putative GBM CTC clusters, we found that they were carrying at total of 116 variants (possibly both driver and passenger mutations) comprised in a total of 58 cancer-associated genes (i.e., genes that were reported in the COSMIC database and whose corresponding substitutions were leading to either structural variants, premature STOP codon, or missense variants) including *APC*, *XPO1*, *TFRC*, *JAK2*, *BRCA2*, *ERBB4* and *ALK*, in addition to the GBM-associated genes *ATM*, *PMS2* and *POLE* (Supplementary Fig. 6c and Suppl. Table 3). Further, despite the notorious difficulties in capturing the entire mutational heterogeneity of an individual's primary GBM with a single biopsy,[11] we asked whether any of the mutations found at the level of GBM CTC clusters were also detectable in the primary tumour. Deep sequencing of a matched primary tumour biopsy of patient #4 revealed that 16 out of 116 GBM CTC cluster mutations were indeed also found in the primary tumour (Suppl. Table 3), including mutations in *POLE*, *BRCA2*, *ALK*, and *TFRC* (Suppl. Table 4). Importantly, several of these loci that contained mutations both in the primary tumour and GBM CTC clusters were covered and found to contain a wild-type sequence in the WBC pool control sample (Suppl. Table 4), conclusively demonstrating the GBM origin of the isolated CTC clusters.

Together, our results provide the first evidence that GBM can release clustered CTCs, characterised by a distinct mutational profile, which pass beyond the blood–brain barrier and reach the peripheral circulation.

DISCUSSION

Although CTC clusters have been highlighted as highly-efficient metastatic precursors,[9] they remain a poorly characterised feature of the metastatic process, mainly due to their dilution factor in the blood of patients and challenges to isolate them without disrupting cell–cell interactions. As further confounding factors, physiological features such as the localisation of the primary tumour, circulation dynamics and entrapment in capillary beds may strongly influence the number of detected CTC clusters in the peripheral circulation of a given patient.

Our work provides the first evidence that circulating glioblastoma clusters can overcome the blood–brain barrier and reach the peripheral circulation. Yet, much remains to be understood in regard to their biology and clinical relevance in this disease, possibly using larger patient cohorts. In the future, it will be of interest to explore the transcriptome of CTC clusters, to define their genetic heterogeneity as compared with the primary tumour, and to assess whether or not their presence in the peripheral circulation correlates with GBM aggressiveness. We speculate that since cancer cells are able to enter the blood circulation as clusters in a non-epithelial cancer such as GBM, a limited cell–cell junction repertoire[12–14] may still be sufficient to sustain CTC clustering and ensure the preservation of multicellular structures beyond vascular barriers. Further studies will be needed to test this hypothesis and to identify key factors involved in CTC cluster generation and passage through physical barriers. In addition, the identification of the essential cell–cell junction requirements to maintain CTC clustering may lead to the development of novel cluster-targeting agents.

ACKNOWLEDGEMENTS

We thank all patients that participated in the study, all involved clinicians and study nurses, as well as all members of the Aceto lab for feedback and discussions. We thank Dr Nina Tunariu (The Royal Marsden Hospital) and Dr Nick Brown (University College London) for support. We thank Dr Christian Beisel and all members of the Genomics Facility Basel (ETH Zürich) for sequencing support. We also thank Basilea Pharmaceutica International Ltd. for financial support, trial setup and organisation, samples withdrawal, scheduling and shipment.

AUTHOR CONTRIBUTIONS

I.K., F.C.-G., M.E. and N.A. designed the study, performed the experiments and drafted the manuscript. S.G., B.S. and R.S. performed immunofluorescence staining and mouse experiments. M.M., N.C., S.C., F.B., S.A., H.L., T.R.J.E., E.P., R.K. and J.L. identified and collected patient samples and analysed the clinical data.

ADDITIONAL INFORMATION

Supplementary Information is available for this paper at https://doi.org/10.1038/s41416-018-0186-7.

Competing interests: The authors declare no competing interests.

Funding: Research in the Aceto lab is supported by the European Research Council, the Swiss National Science Foundation, the Swiss Cancer League, the Basel Cancer League, Basilea Pharmaceutica International Ltd., the two Cantons of Basel through the ETH Zürich, and the University of Basel.

Ethics approval and consent to participate: Patients were participants in the GBM arm of the ongoing study 'An open-label Phase 1/2a study of oral BAL101553 in adult patients with advanced solid tumours and in adult patients with recurrent or progressive glioblastoma or high-grade glioma' (NCT02490800). The study was conducted in accordance with the Declaration of Helsinki (2000) and the International Conference on Harmonisation Guidelines for Good Clinical Practice. The study was approved by each study centre's Research Ethics Committee and all patients provided written informed consent prior to enrollment. Mouse experiments were carried out according to institutional and cantonal guidelines (approved mouse protocol 2781, cantonal veterinary office of Basel-City).Consent for publicationAll patients enroled in the study have signed a written informed consent before enrolment and agreed to data publication in anonymised form.**A**vailability of data and materialExome-sequencing data have been deposited to the European Nucleotide Archive (ENA, EMBL-EBI) with the accession number PRJEB27438. Competing interests: The authors declare no competing interests.

REFERENCES

1. Omuro, A. & DeAngelis, L. M. Glioblastoma and other malignant gliomas: a clinical review. *JAMA* **310**, 1842–1850 (2013).
2. Beauchesne, P. Extra-neural metastases of malignant gliomas: myth or reality? *Cancers (Basel)* **3**, 461–477 (2011).
3. Hamilton, J. D. et al. Glioblastoma multiforme metastasis outside the CNS: three case reports and possible mechanisms of escape. *J. Clin. Oncol.* **32**, e80–e84 (2014).
4. Gkountela, S., Szczerba, B., Donato, C. & Aceto, N. Recent advances in the biology of human circulating tumour cells and metastasis. *ESMO Open* **1**, e000078 (2016).
5. Friedmann-Morvinski, D. Glioblastoma heterogeneity and cancer cell plasticity. *Crit. Rev. Oncog.* **19**, 327–336 (2014).
6. Macarthur, K. M. et al. Detection of brain tumor cells in the peripheral blood by a telomerase promoter-based assay. *Cancer Res.* **74**, 2152–2159 (2014).
7. Muller, C. et al. Hematogenous dissemination of glioblastoma multiforme. *Sci. Transl. Med.* **6**, 247ra101 (2014).
8. Sullivan, J. P. et al. Brain tumor cells in circulation are enriched for mesenchymal gene expression. *Cancer Discov.* **4**, 1299–1309 (2014).
9. Aceto, N. et al. Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis. *Cell* **158**, 1110–1122 (2014).
10. Macaulay, I. C., Ponting, C. P. & Voet, T. Single-cell multiomics: multiple measurements from single cells. *Trends Genet.* **33**, 155–168 (2017).
11. Sottoriva, A. et al. Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. *Proc. Natl Acad. Sci. USA* **110**, 4009–4014 (2013).
12. Grundy, T. J. et al. Differential response of patient-derived primary glioblastoma cells to environmental stiffness. *Sci. Rep.* **6**, 23353 (2016).
13. Hitomi, M. et al. Differential connexin function enhances self-renewal in glioblastoma. *Cell Rep.* **11**, 1031–1042 (2015).
14. Paul, C. D., Mistriotis, P. & Konstantopoulos, K. Cancer cell motility: lessons from migration in confined spaces. *Nat. Rev. Cancer* **17**, 131–140 (2017).

ESMO Open Cancer Horizons

Recent advances in the biology of human circulating tumour cells and metastasis

Sofia Gkountela, Barbara Szczerba, Cinzia Donato, Nicola Aceto

To cite: Gkountela S, Szczerba B, Donato C, *et al*. Recent advances in the biology of human circulating tumour cells and metastasis. *ESMO Open* 2016;**1**:e000078. doi:10.1136/esmoopen-2016-000078

► Prepublication history for this paper is available online. To view these files please visit the journal online (http://dx.doi.org/10.1136/esmoopen-2016-000078).

Received 25 May 2016
Revised 29 June 2016
Accepted 30 June 2016

Cancer Metastasis, Department of Biomedicine, University of Basel, Basel, Switzerland

Correspondence to
Professor Nicola Aceto;
Nicola.Aceto@unibas.ch

ABSTRACT

The development of a metastatic disease is recognised as the cause of death of over 90% of patients diagnosed with cancer. Understanding the biological features of metastasis has been hampered for a long time by the difficulties to study widespread cancerous lesions in patients, and by the absence of reliable methods to isolate viable metastatic cells during disease progression. These difficulties negatively impact on our ability to develop new agents that are tailored to block the spread of cancer. Yet, recent advances in specialised devices for the isolation of circulating tumour cells (CTCs), hand-in-hand with technologies that enable single cell resolution interrogation of their genome and transcriptome, are now paving the way to understanding those molecular mechanisms that drive the formation of metastasis. In this review, we aim to summarise some of the latest discoveries in CTC biology in the context of several types of cancer, and to highlight those findings that have a potential to improve the clinical management of patients with metastatic cancer.

INTRODUCTION

Despite remarkable improvements in early detection of cancerous lesions, combined with surgical removal and treatment of the primary tumour, the development of a metastatic disease remains the main cause of death for the vast majority of patients with cancer. Currently, worldwide, more than seven million people per year die as a consequence of a metastatic disease (WHO). Clearly, these numbers reflect our limited understanding of the biology of cancer in the metastatic setting, and the need to overcome a number of clinical as well as technical limitations to better comprehend how to efficiently target metastasis in patients.

Clinically, patients with a metastatic disease are often characterised by the presence of one or more micrometastatic and macrometastatic foci throughout various organs.[1] Only some of these foci are clinically detectable at the time of metastasis diagnosis, while many more foci are likely to be present but below detection limit, yet posing a significant risk in terms of disease progression.[1 2] In this context, treatment strategies for patients with metastasis are often based on the pathological and molecular characterisation of the primary tumour, while little or no information is available from the various metastatic lesions that are present at that specific moment. A main reason for this is that metastases are hardly accessible for direct sampling. However, the consequence is that we still treat metastasis based on information obtained from a primary tumour, an approach that has led to no success. Further, we now started to be aware of the fact that metastasis is an evolutionary process, where metastatic subclones with a unique mutational profile may emerge along with cancer progression at any time, resulting in a high degree of heterogeneity and significant differences from the primary tumour of origin.[3]

Technically, our current understanding of cancer heterogeneity, especially in the metastatic setting, argues that even if we were able to biopsy most of the metastatic lesion in a given patient, we would still face the issue of cellular heterogeneity within each lesion.[4] However, newly established protocols now enable a single cell resolution interrogation of the genome and transcriptome of cancer cells, and their application to the metastasis field holds the promise to define its molecular drivers with high precision.[5]

While metastasis remains very challenging for a direct biopsy, recent developments in microfluidics technologies are enabling the capture of live circulating tumour cells (CTCs) from the blood of patients with various types of cancer.[6] Most excitingly, given the short half-life of CTCs in circulation,[7] in the metastatic setting CTCs are a direct derivative of metastasis. Pragmatically, they can be seen as an opportunity to isolate, in real time, live cancer cells that

are derived from proliferating metastatic lesions in patients, thereby empowering a single cell resolution analysis of metastasis from a minimally invasive liquid biopsy.

It is important to mention that CTCs are extraordinarily rare in the blood of patients with cancer (on average, one CTC per billion normal blood cells), even in those patients with progressing metastatic disease.[8] In fact, detection of CTCs greatly varies depending on the technology used for their isolation. The majority of platforms currently available for CTC isolation rely on the expression of cell surface markers or physical properties to distinguish CTCs from normal blood components.[9–15] For epithelial cancer types that generally express high levels of epithelial cell adhesion molecule (EpCAM), the CellSearch system is currently the only Food and Drug Administration (FDA)-approved device in the clinical setting.[12 13 16] On the other hand, size-based enrichment platforms, such as Parsortix and ScreenCell, take advantage of the slightly bigger size of CTCs compared with red and white blood cells (WBCs; ~12–25 µm for a single CTC vs 8 and 7–15 µm for a red and WBC, respectively).[17–19] Key to CTC enumeration and characterisation are also newly developed microfluidic devices, such as the spiral biochip based on hydrodynamic forces, or the CTC-iChip which uses a combination of hydrodynamic cell separation and immunomagnetic depletion of antibody-tagged WBCs to isolate larger CTCs.[15 20] However, an unbiased assessment of CTC number and molecular characteristics still requires to overcome a very high degree of technical difficulties, reviewed elsewhere, such as the need for antigen-independent CTC enrichment techniques,[14] devices for single cell manipulation with minimal losses,[21 22] DNA and RNA amplification protocols for single cells with low sequence biases[23] and bioinformatics tools that facilitate single cell data normalisation and analysis.[24] These technologies are now within reach, and are likely to enable a comprehensive characterisation of metastatic cells in the near future.

In this review, we aim to summarise some of the most significant discoveries in CTC biology in different cancer types, focusing on those that are likely to impact our understanding of the metastatic process.

THE BIOLOGY OF CTCS IN DIFFERENT CANCER TYPES

The analysis of CTCS in several cancer types has already allowed a better understanding of the metastatic process, as well as it has highlighted the use of CTCs as a non-invasive source of information for individualised medicine. However, much work remains to be done to gain those insights that will allow the development of new metastasis-tailored therapies. While most of our understanding of CTC biology derives from the analysis of breast and prostate CTCs, recently we also witnessed important advances in colorectal, pancreatic and lung cancer, as well as in melanoma and glioblastoma multiforme (GBM).

Breast cancer

Breast cancer is the second most common cancer in the world, and by far the most common cancer in women, with more than 1.6 million new cases each year. A vast proportion of breast cancers are cured with surgery; however, the development of metastasis still accounts for more than 500 000 deaths per year worldwide (WHO).

Breast cancer is the cancer in which most studies related to CTCs have been carried out. For example, breast CTCs have been shown to be predictors of decreased survival in patients with early breast cancer, before or after chemotherapy, with prognosis being worst in those patients with at least five CTCs per 30 mL of blood.[25] Along these lines, systemic spread of breast cancer has been shown to start early during tumour progression in patients and mouse models, where CTCs are released from a growing primary tumour mass and disseminate to distant organs, leading to the development of metastasis.[26] In the metastatic setting, high CTC counts have also been shown to be predictive of bad prognosis, including those patients who were newly diagnosed with metastatic breast cancer and were about to start first-line systemic treatment.[27 28]

One of the key aspects that emerged from the analysis of breast CTCs is their remarkable heterogeneity, both considering the expression of specific cancer-associated markers, and also their phenotypic characteristics such as tumour-seeding potential. For instance, CTCs frequently lack estrogen receptor (ER) expression in patients with metastatic breast cancer who were diagnosed with ER-positive primary tumours, and these CTCs show a high degree of intrapatient heterogeneity which may reflect a mechanism to escape endocrine therapy.[29 30] Similarly, the expression of the human epidermal growth factor receptor 2 (HER2) on CTCs is often not concordant with the HER2 status of the primary tumour, and a subset of patients with HER2-negative primary tumour will develop HER2-positive CTCs during disease progression.[31] These findings have clear implications for what concerns targeted therapy approaches, and highlight the need to characterise CTCs in real time to define the best treatment for individual patients.

In terms of phenotypic characteristics, significant efforts have been conducted on breast CTCs to identify mesenchymal-like cells among them, and also determine which CTCs in a patient are more likely to successfully form metastasis at a distant site. Epithelial-to-mesenchymal transition (EMT) has been observed in human CTCs from patients with breast cancer, highlighting the occurrence of mesenchymal-like CTCs during disease progression.[8] In an index patient who received longitudinal blood monitoring for epithelial and mesenchymal CTCs along treatment, reversible shifts between these cell fates accompanied each cycle of response to therapy and disease progression, suggesting that EMT may occur as a

consequence to treatment failure.[8] Although EMT has now been shown to occur in human specimens, it is important to highlight that the requirement of EMT for the development of metastasis is still highly controversial.[32 33]

A clinically relevant question regarding CTC biology is whether dissecting CTC heterogeneity in breast cancer can help to identify CTC populations with increased tumour-seeding potential. In this regard, it was recently shown that in patients with breast cancer, the metastasis initiating CTCs were confined in a subpopulation of EpCAM-positive CTCs that also expressed CD44, CD47 and the tyrosine kinase receptor c-MET.[34] This finding was supported by evidence in mouse xenograft models, and based on the increased ability of $EpCAM^+/CD44^+/CD47^+/c\text{-}MET^+$ CTCs to form bone, lung and liver metastases when transplanted in the femoral cavity of immunocompromised mice, compared with the entire population of EpCAM-expressing CTCs.[34] These metastasis-initiating markers were also expressed in patient metastases as judged by histological assessment.[34] Interestingly, another recent study identified a brain metastasis expression signature in the EpCAM-negative fraction of breast CTCs, defined as a subpopulation of $HER2^+/EGFR^+/HPSE^+/Notch1^+$ CTCs.[35] These results indicate that multiple subpopulations of metastasis-initiating CTCs can be present in the peripheral blood of patients with breast cancer.

In addition to markers, certain physical properties of CTCs were also recently linked to increased metastatic potential. In mouse xenograft models, CTC clusters were shown to carry a 23-fold to 50-fold increased metastatic potential compared with single CTCs.[7] In patients with breast cancer, the presence of CTC clusters is also associated with decreased metastasis-free survival and the development of new metastatic foci.[7] Interestingly, these CTC clusters were shown to arise from oligoclonal expansion of tumour cell groupings, rather than from the aggregation of single CTCs in the vasculature or the proliferation of a single CTC. Detailed molecular profiling of single and clustered CTCs from patients with metastatic breast cancer identified the cell–cell junction component plakoglobin to be required for CTC cluster formation. Further, plakoglobin expression in the primary tumour of patients with breast cancer is associated with a reduced metastasis-free survival.[7] In a separate study, CTC clusters were also shown to have an increased metastatic potential and to express keratin14, a marker that was previously found expressed in basal breast cells.[36] Additionally, CTC clusters are found in some instances in association to WBCs, platelets and fibroblasts, which in turn may promote survival and shape the molecular profile of the cells they are in contact with, further promoting heterogeneity and tumour-seeding potential.[7 37–39] These results highlight CTC clusters as extraordinarily efficient metastatic precursors in breast cancer, and the need to further refine their characterisation in breast and other cancer types, to identify their key vulnerabilities.

Prostate cancer

Prostate cancer is the second most common cancer in men, with ~1.1 million new cases each year worldwide, leading to more than 300 000 deaths per year (WHO). Patients with prostate cancer succumb to this disease because of the development of a metastatic disease that is resistant to therapeutic agents, and that usually involves the bone as the primary metastatic site.[40]

In advanced prostate cancer, similarly to breast cancer, CTC enumeration has been used as a prognostic tool for disease progression and survival;[41] however, in localised prostate cancer, a clear correlation between CTC numbers and clinical outcome has not been found.[42 43]

More generally, prostate cancer is frequently responsive to androgen deprivation therapy, given the high expression and requirement of the androgen receptor (AR) in the development and progression of this disease.[40] However, the effectiveness of AR inhibitors in recurrent metastatic disease is highly variable. While direct sampling of metastatic lesions in these patients is challenging, CTC analysis may be key to reveal the mechanisms of innate or acquired resistance to AR-targeted therapies. For this reason, most recent studies have been focusing on performing detailed molecular characterisation of CTCs isolated from patients with metastatic prostate cancer. For example, whole exome sequencing of CTCs was employed to detect somatic single nucleotide variants in patients with prostate cancer.[44] In this study, the mutational profile of CTCs was compared with that of the primary tumour and metastasis and showed a high degree of similarity, establishing a proof-of-concept that CTC genomics can be used in the clinic as a non-invasive method to assess the mutational landscape of metastatic prostate cancer.[44] In a similar approach, whole genome amplification of CTCs isolated from castration-resistant patients with prostate cancer was used to assess copy number aberrations.[45] In this study, the majority of aberrations found in CTCs were also present in the primary tumour; however, copy number gains at the AR locus were found specifically in CTCs, arguing that they could have emerged as a consequence of AR therapy resistance.[45]

In a recent study that employed single cell resolution RNA sequencing of CTCs, putative stem cell markers, such as ALDH7A1, CD44 and KLF4, as well as markers for cell proliferation were enriched in prostate CTCs compared with primary tumours.[46] Of note, in this study, single prostate CTCs displayed a significant heterogeneity, including the expression of AR gene mutations and splicing variants, such as AR-v7, previously shown to confer resistance to antiandrogen therapies.[46 47] Retrospective analysis of CTCs from patients who were progressing under AR-targeted treatment, compared with untreated cases, showed a remarkable activation of non-canonical Wnt signalling pathway. Ectopic expression of Wnt5a in prostate cancer cells attenuated the antiproliferative effects of AR inhibition, whereas its suppression in drug-resistant cells restored partial sensitivity

to anti-AR treatment.[46] Altogether, these studies provide evidence that CTC analysis in patients with metastatic prostate cancer is an opportunity to reveal those key mechanisms of resistance to AR inhibitors, as well as to identify those patients who would benefit the most from a targeted treatment.

Colorectal cancer

Colorectal cancer (CRC) is the second most common cancer in women and the third most common cancer in men, with a total of more than 1.3 million new cases each year. Owing to the development of a metastatic disease, usually affecting the liver at first, almost 700 000 patients succumb to this disease yearly worldwide (WHO).

CTCs have been observed in patients with CRC, in the metastatic as well as non-metastatic setting.[48–50] Notably, the number of CTCs in CRC was also shown to be important in predicting the development of tumour recurrence and emergence of distant metastasis.[48 51–53] More generally however, the number of CTCs detected in the peripheral blood of patients with CRC is much lower compared with other cancers such as breast cancer. Along this line, a recent study compared CTC numbers in blood drawn from peripheral or mesenteric blood in patients with CRC, and found a higher number of CTCs in mesenteric blood samples (median of 2.7–4 CTCs in mesenteric blood vs 0–2 CTCs in peripheral blood), indicating that a considerable portion of CTCs are likely to be trapped in the liver before they reach the peripheral circulation, as the liver is the first filter organ that CTC will encounter on release from the primary tumour mass in the colon.[54] Thus, CTC isolation efforts should also take into account the site of the primary tumour mass (and eventually of each of the metastatic deposits) as well as the blood circulation dynamics to efficiently capture the most viable cancer cells in circulation.

CTCs from patients with metastatic CRC have also been the first to be sequenced at the single cell level. Single cell array comparative genomic hybridization (CGH) and parallel sequencing of 68 CRC-associated genes confirmed the presence of driver mutations in genes such as APC, KRAS or PIK3CA in the primary tumour, metastasis and corresponding CTCs from two patients.[55] However, certain mutations were only visible in CTCs, probably due to their low frequency in the primary tumour and metastatic deposits.[55] These results suggest that liquid biopsy in patients with CRC is highly promising strategy to monitor tumour genomes in real time and facilitate personalised therapy.

Pancreatic cancer

Pancreatic ductal adenocarcinoma (PDAC) is diagnosed in more than 300 000 individuals each year (World Cancer Research Fund, WCRF). Owing to its highly aggressive nature and the fact that early stages of this cancer do not usually produce symptoms, PDAC is almost always fatal, with a 5-year survival rate at ~5% (WCRF). The development of metastasis is the main cause of death in patients with this disease.

Patients with PDAC have been examined for the presence of CTCs in their bloodstream, which were detected with various technologies and at various concentrations.[56–59] However, individually, these studies failed to find a clear correlation between CTC abundance and prognosis. Taken together into a meta-analysis on 623 patients, CTC-positive patients with PDAC showed a worse prognosis compared with patients with no detectable CTCs, independently from the detection method that was used.[60]

Alterations in the gene encoding KRAS are a key feature in PDAC, and KRAS mutations have been accordingly found in PDAC CTCs.[61] Along the line of a molecular characterisation of PDAC CTCs, single cell resolution RNA sequencing of human and mouse PDAC CTCs has highlighted the expression of extracellular matrix (ECM) genes in these cells,[59] as well as the activation of Wnt signalling.[62] In these studies, mouse PDAC CTCs showed upregulation of Wnt2,[62] low-proliferative signatures, enrichment of the stem cell-associated gene ALDH1A2, a biphenotypic expression of epithelial and mesenchymal markers and the expression of IGFBP5, a gene transcript enriched at the epithelial–stromal interface in PDAC primary tumours.[59] Yet, both mouse and human PDAC CTCs displayed a very high expression of stromal-derived ECM proteins, including SPARC, whose knockdown in cancer cells suppressed cell migration and invasiveness.[59] These results highlight that PDAC CTCs may employ expression of Wnt signalling effectors and ECM proteins to facilitate their *route* to metastasis. However, further studies are required to identify effective therapeutic targets with potential to suppress the spread of PDAC cells.

Lung cancer

Lung cancer has been the most common cancer in the world for several decades, with ~1.8 million new cases diagnosed each year (WHO). Lung cancer is also the most common cause of cancer-related death, with ~1.6 million deaths worldwide due to this disease, yearly.

Lung cancer is a key example in targeted therapy approaches, since patients with non-small-cell lung cancer (NSCLC) harbouring activating mutations in the EGFR gene demonstrate a significant progression-free survival benefit when treated with EGFR tyrosine kinase inhibitors (TKIs).[63] However, the majority of patients that are initially responding will develop acquired resistance after 12–24 months of treatment. Mechanisms to TKI resistance include the development of a recurrent T790M EGFR mutation, amplification of signalling molecules that bypass EGFR inhibition (such as MET and HER2), mutations in other oncogenic drivers (eg, PIK3CA and B-RAF) and conversion to small-cell lung cancer (SCLC).[64–69] In this context, the possibility to interrogate lung cancer genotype in real time through liquid biopsies is of paramount importance. In

patients with EGFR-mutant NSCLS, it was previously shown that an allele-specific assay was able to detect the emergence of T790M mutations in CTCs during first-line therapy.[70] Subsequently, other studies confirmed that the analysis of lung CTCs can enable the monitoring of evolving tumour genotype in some patients.[71–74] In addition to their genotype, physical characteristics of NSCLC CTCs have been studied, revealing that NSCLC CTCs appear as single or clustered, with the latter being mostly negative for the proliferation marker Ki67.[75]

In SCLS, CTCs have been detected in great numbers and their abundance clearly correlates with a reduced overall survival.[76] More specifically, patients with more than 50 CTCs per 7.5 mL of blood have an overall survival of 5.4 months, compared with patients with <50 CTCs per 7.5 mL of blood, characterised by an overall survival of 11.5 months.[76] CTCs in SCLC are detected as both single CTCs and CTC clusters, with the latter appearing protected from anoikis and with an increased resistance to cytotoxic drugs.[76] Interestingly, CTCs from patients with SCLC have been also recently employed for transplantation in immunocompromised mice, thus recapitulating the features of the tumour growing in the donor patient.[77] In this model, genomic analysis of the CTC-derived xenografts revealed a high degree of similarity to the original tumour, and a similar responsiveness to platinum and etoposide chemotherapy,[77] thereby providing an excellent platform to guide precision medicine.

Melanoma

Melanoma is diagnosed in more than 230 000 patients per year worldwide, and approximately one-fifth of these patients are lost each year (WCRF). The main cause of death in patients with melanoma is the development of a systemic metastatic disease, affecting most frequently organs such as the liver, bone and brain.[78]

In the past 5 years however, a paradigm shift has occurred in the treatment of this disease. First, a better understanding of the genetic landscape of melanoma has allowed the development of targeted therapies with efficacy against this disease. One above all, is the discovery that B-RAF oncogene is mutated in ~50% of melanomas, and that patients with this genotype benefit from therapy with B-RAF and MEK inhibitors,[79–81] although most will develop resistance within 12 months.[82–84] Second, the understanding of key pathways controlling the immune system has led to the development of immune checkpoint inhibitors such as antibody antagonists of CTLA-4 and PD-1, which individually confer a significant survival benefit to a subset of patients, and even better responses when combined.[85–89] However, at present it is still unclear which patients will benefit from these agents, therefore the identification of biomarkers of response is a priority. More specifically, in melanoma, acquired resistance to therapy seems to be driven by the clonal expansion of resistant tumour cells.[90] While repeated biopsies to study genomic alterations along therapy are invasive, difficult to obtain and prone to be confounded by intratumoural heterogeneity, the analysis of CTCs may result as a powerful weapon to stratify patients in light of the best treatment option.

Since melanoma is not an epithelial cancer, CTCs are extraordinarily hard to isolate from the blood of patients, because they do not express common CTC markers such as EpCAM or epithelial cytokeratins that would distinguish them from normal blood cells. However, their isolation is now achievable through antigen-agnostic techniques.[14] This is enabling a better understanding of their biology, and highlighting the possibility to stratify patients before treatment. Melanoma CTC xenografts have been also recently developed, and hold the promise to serve as a platform to screen various therapeutic agents in vivo, while gaining insights into tumour evolution dynamics.[91] Further, melanoma CTCs have been detected in patients along B-RAF-targeted therapy, with their number increasing during disease progression.[92] Lastly, a recent report has shown that the number of melanoma CTCs in patients is much higher when blood is taken from the arterial, rather than venous, circulation.[93] Along the lines of what is discussed above for CRC, these results indicate a better, although rather inconvenient for the patient, source of blood sampling for melanoma CTCs. Altogether, these studies demonstrate the ability to isolate and characterise melanoma CTCs from patients, and may be key in the future to achieve patient stratification before the administration of targeted therapy or immunotherapy.

Glioblastoma multiforme

Worldwide, there are an estimated 240 000 cases of brain and nervous system tumours per year, with GBM being the most common, and the most lethal, of these tumours (WHO). Unlike other tumours, patients with GBM die because of the consequences of tumour growth in the primary tumour site, and development of metastasis is extremely rare.[94]

Major challenges in the treatment of GBM include the inability to excise tumour cells infiltrating into normal brain tissue, the poor penetration of therapeutic agents into the central nervous system (CNS), the difficulty in distinguishing tumour responses from recurrence using standard imaging criteria and the inherent risks associated with brain biopsies needed to monitor tumour evolution during disease progression.[94] In this context, it has been clear for long time that the possibility to isolate GBM CTCs from liquid biopsies may significantly help in understanding GBM biology. However, until very recently, it has been unclear whether GBM cancer cells would be able to cross the blood–brain barrier and be detectable in the peripheral circulation. Further, similarly to melanoma and unlike many epithelial cancer types, CNS malignancies do not express EpCAM, a marker that is commonly used for CTC detection. Thus, the isolation of GBM CTCs has been hampered by a number of exceptionally hard challenges.

However, in 2014, three groups reported the successful isolation and characterisation of CTCs from the peripheral blood of patients with GBM.[95–97] In a first study, glial fibrillary acidic protein-expressing CTCs were detected in 29/141 patients with GBM.[95] These CTCs were identified in the density gradient purified mononuclear fraction of peripheral blood and further validation was based on the expression of EGFR mutations and aberrations that matched the primary GBM tumours. However, there was no correlation between CTC enumeration and clinical outcome before or after surgery in these patients.[95] In a second study, using a strategy based on telomerase activity, GBM CTCs were successfully detected in a number of patients.[96] Specifically, CTCs were detected in 8/11 preradiotherapy patients as opposed to 1/8 in postradiotherapy patients, indicating that CTC enumeration in GBM could be useful in identifying patients who are at high risk of recurrence.[96] In the third study, using the CTC-iChip platform combined with a specific staining optimised to distinguish GBM cells from any other blood cell, CTCs were found in 13/33 patients with GBM.[97] These CTCs were identified using a cocktail of fluorescent probes targeting five known high-grade glioma markers, termed 'STEAM' (SOX2, Tubulin β-3, EGFR, A2B5, c-MET).[97] As further validation, CTCs were shown to harbour EGFR gene amplifications that corresponded to the primary GBM tumour. Additional single cell resolution expression analyses identified a high enrichment of mesenchymal-associated transcripts, such as SERPINE1, TGFB1, TGFB2 and vimentin at the expense of neural lineage markers, compared with matched primary tumours. Interestingly, these mesenchymal markers were also expressed by RNA-ISH at distinct areas of the primary tumour and predominantly at the invasive edge of the deep white matter tracts, the area of the tumour that is associated with GBM cell invasion.[97] Nevertheless, in this study there was no clear correlation between CTC presence and clinical outcome of these patients.[97] Altogether, these studies were instrumental to demonstrate for the first time the presence of CTCs in patients with GBM, and warrant further investigations to gain more insights into the biology of this disease.

Culturing CTCs for personalised medicine

While the analysis of freshly isolated CTCs might be a phenomenal opportunity to stratify patients and to guide precision medicine in the future, the extremely low abundance of these cells in the peripheral blood of patients with cancer remains a challenge in the context of personalised drug screenings. The possibility of expanding CTCs in culture has only very recently been achieved, carrying important implications for personalised medicine.

The first study reporting successful culture of CTCs was performed with samples from patients affected by brain metastatic breast cancer.[35] In this study, a fraction of EpCAM-negative CTCs was found to carry a HER2/EGFR/HPSE/Notch1 protein signature and to be particularly prone to form brain metastasis. These cells were cultured, and on transplantation in mice, these CTC cell lines were highly invasive and capable to generate brain and lung metastasis in animal models.[35] However, the first example of CTC cultures with the aim of individualised drug testing was provided in a different study.[98] In that study oligoclonal CTC cultures were derived from six patients with ER-positive breast cancer, and subjected to genome sequencing of a panel of cancer-associated mutational hotspots. Data analysis revealed pre-existing as well as acquired mutations in PIK3CA, ESR1 and FGFR2 genes, among others. Drug sensitivity testing ex vivo and xenografts of each CTC cell line revealed key (personalised) vulnerabilities as a proof-of-concept.[98] In another study, CTCs as well as tumour biopsies derived from patients with prostate cancer were expanded as long-term organoid cultures.[99] Seven newly generated organoid cell lines were shown to recapitulate the molecular diversity of prostate cancer subtypes, including TMPRSS2-ERG fusion, SPOP, FOXA1 and PIK3R1 mutations, SPINK1 overexpression, and loss of CHD1, p53 and RB tumour suppressor.[99] Other studies that followed could then show for the first time a successful CTC culture establishment from CRC[100] and lung cancer,[101] paving the way to a detailed molecular and phenotypic analysis of CTCs in these diseases as well.

Altogether, several groups have now successfully established long-term CTC cultures from different cancer types. In the context of personalised medicine however, much work remains to be done. For instance, establishment of CTC-derived cell lines nowadays still requires several months, and it is only possible from a restricted number of patients, usually those with the highest numbers of CTCs.[98–101] During this time, most CTCs isolated from a patient will die in culture, and only some will be able to successfully grow and establish a cell line. During this process, the corresponding patient in the clinic is likely to undergo additional treatment cycles, which are expected to reshape the molecular portrait of his/her disease.[102 103] In this scenario, a drug screening on CTC-derived cells would not be up to date with the patient's disease. For CTC cultures to become a strategy that enables real-time personalised medicine, much progress needs to be made in order to achieve drug susceptibility testing within only a few weeks, if not just days, after the blood is drawn. Thus, increasing the success rates of CTC culture assays along with the development of more rapid culture strategies is of paramount importance for achieving personalised drug screenings from liquid biopsy, as well as to enable most patients with cancer to benefit from this approach.

CONCLUSIONS

In the past few years, the CTC field has witnessed outstanding advances. First, it is now possible to efficiently

isolate CTCs in an antigen-agnostic fashion.[7 15 18 20] This allows an unbiased CTC enrichment strategy in epithelial cancers, and also permits the isolation of CTCs from cancers of non-epithelial origin, such as GBM and melanoma. Second, several CTC-enrichment technologies are now able to release viable CTCs in solution, thereby empowering their micromanipulation or culture, and separation from contaminant blood cells after first-step enrichment.[7 46 98] Third, single cell resolution sequencing of the genome or transcriptome of CTCs has been achieved,[5] showing that it may represent an extraordinary opportunity to characterise the mutational profile of metastatic cells in real time, to interrogate patient samples longitudinally during treatment, as well as to dissect fundamental pathways that orchestrate the metastatic process. Fourth, several groups have shown the ability to expand CTCs in culture or as xenografts with the goal of testing individualised drug susceptibility, and creating new CTC-derived lines that represent highly clinically relevant models to study how metastasis occurs at the molecular level.[98–101]

These extraordinary discoveries in the CTC field, however, should be seen as the starting point of a journey that promises to bring liquid biopsies into clinical practice. Significant steps ahead are urgently needed to achieve standardised protocols for real-time CTC monitoring and molecular interrogation, early during primary tumour onset and also later during metastatic disease progression, most likely in conjunction with the analysis of cell-free DNA in patients with cancer.[104] At the same time, CTC culturing needs to be achieved in a much faster time frame in order to benefit patients. Last but not least, several outstanding questions remain unanswered in the CTC and metastasis fields. For example, we still do not know what triggers the generation of CTCs (single or clustered) from a primary tumour or metastatic deposit, what is the true evolution pattern of metastasis at the single cell level before and after therapy and what are the targets to inhibit in order to prevent or suppress the haematogenous spread of cancer cells in patients. Answers to these questions are now within reach, and hold great promise to improve the clinical management of patients who suffer from metastatic cancers.

Acknowledgements The authors apologise to their colleagues whose work could not be cited owing to space constraints. The authors thank the European Research Council (ERC), the Swiss National Science Foundation (SNSF), the Krebsliga Beider Basel, the L. & Th. La Roche Foundation and the University of Basel for supporting the work. They are also grateful to all members of NA's laboratory for critical review of the manuscript.

Contributors SG, BS, CD and NA reviewed the literature and wrote the manuscript. All authors have read and approved the final version of the manuscript.

Funding Swiss National Science Foundation (PP00P3_163938); European Research Council (678834); Universität Basel; Krebsliga Beider Basel.

Competing interests NA is an inventor in patent WO2015061091 on 'Treating cancer, by measuring level of CTC clusters in sample obtained from patient with breast or epithelial cancer, administering treatment to prevent or reduce metastasis' owned by the Massachusetts General Hospital, Boston, Massachusetts, USA.

Provenance and peer review Commissioned; externally peer reviewed.

REFERENCES

1. Valastyan S, Weinberg RA. Tumor metastasis: molecular insights and evolving paradigms. *Cell* 2011;147:275–92.
2. Klein CA. Parallel progression of primary tumours and metastases. *Nat Rev Cancer* 2009;9:302–12.
3. Turajlic S, Swanton C. Metastasis as an evolutionary process. *Science* 2016;352:169–75.
4. Navin N, Kendall J, Troge J, *et al*. Tumour evolution inferred by single-cell sequencing. *Nature* 2011;472:90–4.
5. Navin NE. The first five years of single-cell cancer genomics and beyond. *Genome Res* 2015;25:1499–507.
6. Alix-Panabières C, Pantel K. Circulating tumor cells: liquid biopsy of cancer. *Clin Chem* 2013;59:110–18.
7. Aceto N, Bardia A, Miyamoto DT, *et al*. Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis. *Cell* 2014;158:1110–22.
8. Yu M, Stott S, Toner M, *et al*. Circulating tumor cells: approaches to isolation and characterization. *J Cell Biol* 2011;192:373–82.
9. Stott SL, Hsu CH, Tsukrov DI, *et al*. Isolation of circulating tumor cells using a microvortex-generating herringbone-chip. *Proc Natl Acad Sci USA* 2010;107:18392–7.
10. Vona G, Sabile A, Louha M, *et al*. Isolation by size of epithelial tumor cells: a new method for the immunomorphological and molecular characterization of circulating tumor cells. *Am J Pathol* 2000;156:57–63.
11. Riahi R, Gogoi P, Sepehri S, *et al*. A novel microchannel-based device to capture and analyze circulating tumor cells (CTCs) of breast cancer. *Int J Oncol* 2014;44:1870–8.
12. Beije N, Jager A, Sleijfer S. Circulating tumor cell enumeration by the CellSearch system: the clinician's guide to breast cancer treatment? *Cancer Treat Rev* 2015;41:144–50.
13. Miller MC, Doyle GV, Terstappen LW. Significance of circulating tumor cells detected by the CellSearch system in patients with metastatic breast colorectal and prostate cancer. *J Oncol* 2010;2010:617421.
14. Aceto N, Toner M, Maheswaran S, *et al*. En route to metastasis: circulating tumor cell clusters and epithelial-to-mesenchymal transition. *Trends Cancer* 2015;1:44–52.
15. Ozkumur E, Shah AM, Ciciliano JC, *et al*. Inertial focusing for tumor antigen-dependent and -independent sorting of rare circulating tumor cells. *Sci Transl Med* 2013;5:179ra47.
16. Went PT, Lugli A, Meier S, *et al*. Frequent EpCam protein expression in human carcinomas. *Hum Pathol* 2004;35:122–8.
17. Zheng S, Lin H, Liu JQ, *et al*. Membrane microfilter device for selective capture, electrolysis and genomic analysis of human circulating tumor cells. *J Chromatogr A* 2007;1162:154–61.
18. Xu L, Mao X, Imrali A, *et al*. Optimization and evaluation of a novel size based circulating tumor cell isolation system. *PLoS ONE* 2015;10:e0138032.
19. Freidin MB, Tay A, Freydina DV, *et al*. An assessment of diagnostic performance of a filter-based antibody-independent peripheral blood circulating tumour cell capture paired with cytomorphologic criteria for the diagnosis of cancer. *Lung Cancer* 2014;85:182–5.
20. Warkiani ME, Khoo BL, Wu L, *et al*. Ultra-fast, label-free isolation of circulating tumor cells from blood using spiral microfluidics. *Nat Protoc* 2016;11:134–48.
21. Haupt S, Grützner J, Thier MC, *et al*. Automated selection and harvesting of pluripotent stem cell colonies. *Biotechnol Appl Biochem* 2012;59:77–87.
22. Peeters DJ, De Laere B, Van den Eynden GG, *et al*. Semiautomated isolation and molecular characterisation of single or highly purified tumour cells from CellSearch enriched blood samples using dielectrophoretic cell sorting. *Br J Cancer* 2013;108:1358–67.
23. Gawad C, Koh W, Quake SR. Single-cell genome sequencing: current state of the science. *Nat Rev Genet* 2016;17:175–88.
24. Ning L, Liu G, Li G, *et al*. Current challenges in the bioinformatics of single cell genomics. *Front Oncol* 2014;4:7.

25. Rack B, Schindlbeck C, Jückstock J, *et al*., SUCCESS Study Group. Circulating tumor cells predict survival in early average-to-high risk breast cancer patients. *J Natl Cancer Inst* 2014;106:pii:dju066.
26. Hüsemann Y, Geigl JB, Schubert F, *et al*. Systemic spread is an early step in breast cancer. *Cancer Cell* 2008;13:58–68.
27. Giuliano M, Giordano A, Jackson S, *et al*. Circulating tumor cells as prognostic and predictive markers in metastatic breast cancer patients receiving first-line systemic treatment. *Breast Cancer Res* 2011;13:R67.
28. Cristofanilli M, Budd GT, Ellis MJ, *et al*. Circulating tumor cells, disease progression, and survival in metastatic breast cancer. *N Engl J Med* 2004;351:781–91.
29. Babayan A, Hannemann J, Spötter J, *et al*. Heterogeneity of estrogen receptor expression in circulating tumor cells from metastatic breast cancer patients. *PLoS ONE* 2013;8:e75038.
30. Paoletti C, Larios JM, Muñiz MC, *et al*. Heterogeneous estrogen receptor expression in circulating tumor cells suggests diverse mechanisms of fulvestrant resistance. *Mol Oncol* 2016;pii.
31. Pestrin M, Bessi S, Galardi F, *et al*. Correlation of HER2 status between primary tumors and corresponding circulating tumor cells in advanced breast cancer patients. *Breast Cancer Res Treat* 2009;118:523–30.
32. Fischer KR, Durrans A, Lee S, *et al*. Epithelial-to-mesenchymal transition is not required for lung metastasis but contributes to chemoresistance. *Nature* 2015;527:472–6.
33. Zheng X, Carstens JL, Kim J, *et al*. Epithelial-to-mesenchymal transition is dispensable for metastasis but induces chemoresistance in pancreatic cancer. *Nature* 2015;527:525–30.
34. Baccelli I, Schneeweiss A, Riethdorf S, *et al*. Identification of a population of blood circulating tumor cells from breast cancer patients that initiates metastasis in a xenograft assay. *Nat Biotechnol* 2013;31:539–44.
35. Zhang L, Ridgway LD, Wetzel MD, *et al*. The identification and characterization of breast cancer CTCs competent for brain metastasis. *Sci Transl Med* 2013;5:180ra48.
36. Cheung KJ, Padmanaban V, Silvestri V, *et al*. Polyclonal breast cancer metastases arise from collective dissemination of keratin 14-expressing tumor cell clusters. *Proc Natl Acad Sci USA* 2016;113:E854–63.
37. Sarioglu AF, Aceto N, Kojic N, *et al*. A microfluidic device for label-free, physical capture of circulating tumor cell clusters. *Nat Methods* 2015;12:685–91.
38. Labelle M, Begum S, Hynes RO. Direct signaling between platelets and cancer cells induces an epithelial-mesenchymal-like transition and promotes metastasis. *Cancer Cell* 2011;20:576–90.
39. Duda DG, Duyverman AM, Kohno M, *et al*. Malignant cells facilitate lung metastasis by bringing their own soil. *Proc Natl Acad Sci USA* 2010;107:21677–82.
40. Watson PA, Arora VK, Sawyers CL. Emerging mechanisms of resistance to androgen receptor inhibitors in prostate cancer. *Nat Rev Cancer* 2015;15:701–11.
41. Danila DC, Heller G, Gignac GA, *et al*. Circulating tumor cell number and prognosis in progressive castration-resistant prostate cancer. *Clin Cancer Res* 2007;13:7053–8.
42. Stott SL, Lee RJ, Nagrath S, *et al*. Isolation and characterization of circulating tumor cells from patients with localized and metastatic prostate cancer. *Sci Transl Med* 2010;2:25ra23.
43. Khurana KK, Grane R, Borden EC, *et al*. Prevalence of circulating tumor cells in localized prostate cancer. *Curr Urol* 2013;7:65–9.
44. Lohr JG, Adalsteinsson VA, Cibulskis K, *et al*. Whole-exome sequencing of circulating tumor cells provides a window into metastatic prostate cancer. *Nat Biotechnol* 2014;32:479–84.
45. Magbanua MJ, Sosa EV, Scott JH, *et al*. Isolation and genomic analysis of circulating tumor cells from castration resistant metastatic prostate cancer. *BMC Cancer* 2012;12:78.
46. Miyamoto DT, Zheng Y, Wittner BS, *et al*. RNA-Seq of single prostate CTCs implicates noncanonical Wnt signaling in antiandrogen resistance. *Science* 2015;349:1351–6.
47. Antonarakis ES, Lu C, Wang H, *et al*. AR-V7 and resistance to enzalutamide and abiraterone in prostate cancer. *N Engl J Med* 2014;371:1028–38.
48. Tsai WS, Chen JS, Shao HJ, *et al*. Circulating tumor cell count correlates with colorectal neoplasm progression and is a prognostic marker for distant metastasis in non-metastatic patients. *Sci Rep* 2016;6:24517.
49. Raimondi C, Nicolazzo C, Gradilone A, *et al*. Circulating tumor cells: exploring intratumor heterogeneity of colorectal cancer. *Cancer Biol Ther* 2014;15:496–503.
50. Gorges TM, Stein A, Quidde J, *et al*. Improved detection of circulating tumor cells in metastatic colorectal cancer by the combination of the CellSearch® system and the AdnaTest®. *PLoS ONE* 2016;11:e0155126.
51. Bork U, Rahbari NN, Schölch S, *et al*. Circulating tumour cells and outcome in non-metastatic colorectal cancer: a prospective study. *Br J Cancer* 2015;112:1306–13.
52. Cohen SJ, Punt CJ, Iannotti N, *et al*. Relationship of circulating tumor cells to tumor response, progression-free survival, and overall survival in patients with metastatic colorectal cancer. *J Clin Oncol* 2008;26:3213–21.
53. Lalmahomed ZS, Mostert B, Onstenk W, *et al*. Prognostic value of circulating tumour cells for early recurrence after resection of colorectal liver metastases. *Br J Cancer* 2015;112:556–61.
54. Denève E, Riethdorf S, Ramos J, *et al*. Capture of viable circulating tumor cells in the liver of colorectal cancer patients. *Clin Chem* 2013;59:1384–92.
55. Heitzer E, Auer M, Gasch C, *et al*. Complex tumor genomes inferred from single circulating tumor cells by array-CGH and next-generation sequencing. *Cancer Res* 2013;73:2965–75.
56. Dotan E, Alpaugh RK, Ruth K, *et al*. Prognostic significance of MUC-1 in circulating tumor cells in patients with metastatic pancreatic adenocarcinoma. *Pancreas* Published Online First: 10 Mar 2016. doi:10.1097/MPA.0000000000000619
57. Nagrath S, Sequist LV, Maheswaran S, *et al*. Isolation of rare circulating tumour cells in cancer patients by microchip technology. *Nature* 2007;450:1235–9.
58. Bidard FC, Huguet F, Louvet C, *et al*. Circulating tumor cells in locally advanced pancreatic adenocarcinoma: the ancillary CirCe 07 study to the LAP 07 trial. *Ann Oncol* 2013;24:2057–61.
59. Ting DT, Wittner BS, Ligorio M, *et al*. Single-cell RNA sequencing identifies extracellular matrix gene expression by pancreatic circulating tumor cells. *Cell Rep* 2014;8:1905–18.
60. Han L, Chen W, Zhao Q. Prognostic value of circulating tumor cells in patients with pancreatic cancer: a meta-analysis. *Tumour Biol* 2014;35:2473–80.
61. Kulemann B, Liss AS, Warshaw AL, *et al*. KRAS mutations in pancreatic circulating tumor cells: a pilot study. *Tumour Biol* 2016;37:7547–54.
62. Yu M, Ting DT, Stott SL, *et al*. RNA sequencing of pancreatic circulating tumour cells implicates WNT signalling in metastasis. *Nature* 2012;487:510–13.
63. Lynch TJ, Bell DW, Sordella R, *et al*. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *N Engl J Med* 2004;350:2129–39.
64. Sequist LV, Waltman BA, Dias-Santagata D, *et al*. Genotypic and histological evolution of lung cancers acquiring resistance to EGFR inhibitors. *Sci Transl Med* 2011;3:75ra26.
65. Yu HA, Arcila ME, Rekhtman N, *et al*. Analysis of tumor specimens at the time of acquired resistance to EGFR-TKI therapy in 155 patients with EGFR-mutant lung cancers. *Clin Cancer Res* 2013;19:2240–7.
66. Kobayashi S, Boggon TJ, Dayaram T, *et al*. EGFR mutation and resistance of non-small-cell lung cancer to gefitinib. *N Engl J Med* 2005;352:786–92.
67. Pao W, Miller VA, Politi KA, *et al*. Acquired resistance of lung adenocarcinomas to gefitinib or erlotinib is associated with a second mutation in the EGFR kinase domain. *PLoS Med* 2005;2:e73.
68. Engelman JA, Zejnullahu K, Mitsudomi T, *et al*. MET amplification leads to gefitinib resistance in lung cancer by activating ERBB3 signaling. *Science* 2007;316:1039–43.
69. Zhang Z, Lee JC, Lin L, *et al*. Activation of the AXL kinase causes resistance to EGFR-targeted therapy in lung cancer. *Nat Genet* 2012;44:852–60.
70. Maheswaran S, Sequist LV, Nagrath S, *et al*. Detection of mutations in EGFR in circulating lung-cancer cells. *N Engl J Med* 2008;359:366–77.
71. Marchetti A, Del Grammastro M, Felicioni L, *et al*. Assessment of EGFR mutations in circulating tumor cell preparations from NSCLC patients by next generation sequencing: toward a real-time liquid biopsy for treatment. *PLoS ONE* 2014;9:e103883.
72. Ni X, Zhuo M, Su Z, *et al*. Reproducible copy number variation patterns among single circulating tumor cells of lung cancer patients. *Proc Natl Acad Sci USA* 2013;110:21083–8.
73. Pailler E, Adam J, Barthélémy A, *et al*. Detection of circulating tumor cells harboring a unique ALK rearrangement in ALK-positive non-small-cell lung cancer. *J Clin Oncol* 2013;31:2273–81.
74. Sundaresan TK, Sequist LV, Heymach JV, *et al*. Detection of T790M, the acquired resistance EGFR mutation, by tumor biopsy versus noninvasive blood-based analyses. *Clin Cancer Res* 2016;22:1103–10.

75. Krebs MG, Hou JM, Sloane R, *et al*. Analysis of circulating tumor cells in patients with non-small cell lung cancer using epithelial marker-dependent and -independent approaches. *J Thorac Oncol* 2012;7:306–15.
76. Hou JM, Krebs MG, Lancashire L, *et al*. Clinical significance and molecular characteristics of circulating tumor cells and circulating tumor microemboli in patients with small-cell lung cancer. *J Clin Oncol* 2012;30:525–32.
77. Hodgkinson CL, Morrow CJ, Li Y, *et al*. Tumorigenicity and genetic profiling of circulating tumor cells in small-cell lung cancer. *Nat Med* 2014;20:897–903.
78. Gray-Schopfer V, Wellbrock C, Marais R. Melanoma biology and new targeted therapy. *Nature* 2007;445:851–7.
79. Chapman PB, Hauschild A, Robert C, *et al*. Improved survival with vemurafenib in melanoma with BRAF V600E mutation. *N Engl J Med* 2011;364:2507–16.
80. Flaherty KT, Infante JR, Daud A, *et al*. Combined BRAF and MEK inhibition in melanoma with BRAF V600 mutations. *N Engl J Med* 2012;367:1694–703.
81. Sosman JA, Kim KB, Schuchter L, *et al*. Survival in BRAF V600-mutant advanced melanoma treated with vemurafenib. *N Engl J Med* 2012;366:707–14.
82. Girotti MR, Pedersen M, Sanchez-Laorden B, *et al*. Inhibiting EGF receptor or SRC family kinase signaling overcomes BRAF inhibitor resistance in melanoma. *Cancer Discov* 2013;3:158–67.
83. Larkin J, Ascierto PA, Dréno B, *et al*. Combined vemurafenib and cobimetinib in BRAF-mutated melanoma. *N Engl J Med* 2014;371:1867–76.
84. Long GV, Stroyakovskiy D, Gogas H, *et al*. Combined BRAF and MEK inhibition versus BRAF inhibition alone in melanoma. *N Engl J Med* 2014;371:1877–88.
85. Girotti MR, Saturno G, Lorigan P, *et al*. No longer an untreatable disease: how targeted and immunotherapies have changed the management of melanoma patients. *Mol Oncol* 2014;8:1140–58.
86. Larkin J, Chiarion-Sileni V, Gonzalez R, *et al*. Combined nivolumab and ipilimumab or monotherapy in untreated melanoma. *N Engl J Med* 2015;373:23–34.
87. Robert C, Schachter J, Long GV, *et al*. Pembrolizumab versus ipilimumab in advanced melanoma. *N Engl J Med* 2015;372:2521–32.
88. Snyder A, Makarov V, Merghoub T, *et al*. Genetic basis for clinical response to CTLA-4 blockade in melanoma. *N Engl J Med* 2014;371:2189–99.
89. Tumeh PC, Harview CL, Yearley JH, *et al*. PD-1 blockade induces responses by inhibiting adaptive immune resistance. *Nature* 2014;515:568–71.
90. Shi H, Hugo W, Kong X, *et al*. Acquired resistance and clonal evolution in melanoma during BRAF inhibitor therapy. *Cancer Discov* 2014;4:80–93.
91. Girotti MR, Gremel G, Lee R, *et al*. Application of sequencing, liquid biopsies, and patient-derived xenografts for personalized medicine in melanoma. *Cancer Discov* 2016;6:286–99.
92. Luo X, Mitra D, Sullivan RJ, *et al*. Isolation and molecular characterization of circulating melanoma cells. *Cell Rep* 2014;7:645–53.
93. Terai M, Mu Z, Eschelman DJ, *et al*. Arterial blood, rather than venous blood, is a better source for circulating melanoma cells. *EBioMedicine* 2015;2:1821–6.
94. Omuro A, DeAngelis LM. Glioblastoma and other malignant gliomas: a clinical review. *JAMA* 2013;310:1842–50.
95. Müller C, Holtschmidt J, Auer M, *et al*. Hematogenous dissemination of glioblastoma multiforme. *Sci Transl Med* 2014;6:247ra101.
96. Macarthur KM, Kao GD, Chandrasekaran S, *et al*. Detection of brain tumor cells in the peripheral blood by a telomerase promoter-based assay. *Cancer Res* 2014;74:2152–9.
97. Sullivan JP, Nahed BV, Madden MW, *et al*. Brain tumor cells in circulation are enriched for mesenchymal gene expression. *Cancer Discov* 2014;4:1299–309.
98. Yu M, Bardia A, Aceto N, *et al*. Cancer therapy. Ex vivo culture of circulating breast tumor cells for individualized testing of drug susceptibility. *Science* 2014;345:216–20.
99. Gao D, Vela I, Sboner A, *et al*. Organoid cultures derived from patients with advanced prostate cancer. *Cell* 2014;159:176–87.
100. Cayrefourcq L, Mazard T, Joosse S, *et al*. Establishment and characterization of a cell line from human circulating colon cancer cells. *Cancer Res* 2015;75:892–901.
101. Zhang Z, Shiratsuchi H, Lin J, *et al*. Expansion of CTCs from early stage lung cancer patients using a microfluidic co-culture model. *Oncotarget* 2014;5:12383–97.
102. Lipinski KA, Barber LJ, Davies MN, *et al*. Cancer evolution and the limits of predictability in precision cancer medicine. *Trends Cancer* 2016;2:49–63.
103. McGranahan N, Swanton C. Biological and therapeutic impact of intratumor heterogeneity in cancer evolution. *Cancer Cell* 2015;27:15–26.
104. Sundaresan TK, Haber DA. Fantastic voyage: the future of cancer diagnostics. *Lancet Oncol* 2015;16:1596–8.