I. CTC Review

Thiele J-A, Bethel K, Králíčková M, Kuhn P, 2017, **Circulating Tumor Cells: Fluid Surrogates of Solid Tumors**. Annu Rev Pathol Mech Dis 12:419–447. doi: 10.1146/annurev-pathol-052016-100256 **(IF_{2017} =15.95)**

Circulating Tumor Cells: Fluid Surrogates of Solid Tumors

J.-A. Thiele,[1] K. Bethel,[2] M. Králíčková,[3] and P. Kuhn[4,5]

[1]Biomedical Center, Faculty of Medicine in Pilsen, Charles University in Prague, 323 00 Pilsen, Czech Republic

[2]Scripps Clinic Medical Group, Scripps Clinic, La Jolla, California 92121

[3]Department of Histology and Embryology, Faculty of Medicine in Pilsen, Charles University in Prague, 301 00 Pilsen, Czech Republic

[4]Bridge Institute, Dornsife College of Letters, Arts and Sciences, University of Southern California, Los Angeles, California 90089; email: pkuhn@usc.edu

[5]Department of Biomedical Engineering, Viterbi School of Engineering, University of Southern California, Los Angeles, California 90089

Annu. Rev. Pathol. Mech. Dis. 2017. 12:419–47

The *Annual Review of Pathology: Mechanisms of Disease* is online at pathol.annualreviews.org

This article's doi:
10.1146/annurev-pathol-052016-100256

Keywords

biomarkers, metastasis, liquid biopsy, precision medicine, fluid biopsy, circulating tumor microemboli

Abstract

Evaluation of circulating tumor cells (CTCs) has demonstrated clinical validity as a prognostic tool based on enumeration, but since the introduction of this tool to the clinic in 2004, further clinical utility and widespread adoption have been limited. However, immense efforts have been undertaken to further the understanding of the mechanisms behind the biology and kinetics of these rare cells, and progress continues toward better applicability in the clinic. This review describes recent advances within the field, with a particular focus on understanding the biological significance of CTCs, and summarizes emerging methods for identifying, isolating, and interrogating the cells that may provide technical advantages allowing for the discovery of more specific clinical applications. Included is an atlas of high-definition images of CTCs from various cancer types, including uncommon CTCs captured only by broadly inclusive nonenrichment techniques.

INTRODUCTION

CTC: circulating tumor cell

Improvements in cancer prevention and control as well as changes in medical practice have contributed to cancer death rates decreasing by 22% over the last two decades. However, cancer still remains the second-leading cause of death in the United States, resulting in more than half a million deaths and 1.6 million newly diagnosed patients yearly (1). Cancer is a heterogeneous disease, with each type linked to its own prognostic outcome. Each type can often be further subdivided into a subtype that has a specific phenotype with targets, like hormone receptors or growth factors, that are expressed and allow for specialized targeted therapy (2). However, the clinical experience in both leukemias and solid tumors of the emergence of resistance, as well as many studies describing discordance between the primary site and metastatic lesions, indicates that these biomarker expressions are not static, but dynamic (3–5). The dynamic phenotype of neoplastic cells poses a challenge for oncologists trying to optimize treatment. Ideally, each time the cancer progresses and a therapy change is considered, a fresh biopsy of tumor tissue would be obtained to be certain that it contains the target(s) against which the next round of therapy could be aimed. However, unlike in the leukemias, for which repeat bone marrow biopsies are frequently obtained when disease seems to change its behavior, in solid tumors repeat biopsies are clinically unfeasible and pose risks for patients. A possible solution to the dilemma is to utilize ever-improving technological breakthroughs to obtain small portions of solid tumors that are easily accessible via blood draw. Thus, interest in the concept of the fluid biopsy is growing. A further advantage of this approach, especially in a patient with multiple metastatic sites, is that the circulating component likely represents cells from multiple locations and may thus provide a more informative sample than a single biopsy of a single metastasis. After dissociation from a solid tumor mass, circulating tumor cells (CTCs) travel through the vasculature as single cells or aggregates and contribute to forming a distant metastasis. Although not every CTC represents a potential future metastasis, many distant metastases are considered to be established by hematogenous spread of these cells, rather than by lymphatic or direct intracavitary spread, which likely occurs by a different mechanism. These rare cells therefore provide a rich source of tissue material and may eventually be analyzed by pathologists alongside other cancer cells recovered from body fluids such as pleural and peritoneal fluid. This review discusses what is known regarding the nature of CTCs and the mechanisms of CTC circulation and focuses on the challenges, opportunities, and new insights. Established and novel methods for CTC identification and characterization are discussed, along with these methods' implications for CTC research and potential translation to clinical practice.

CIRCULATING TUMOR CELLS: BACKGROUND

CTCs are surrogates of a tumor in the bloodstream. They were first described nearly 150 years ago by Thomas Ashworth (6), who noticed cells with an unusual morphology in the blood of a patient who had died from cancer. Ashworth considered these cells' possible tumor origin because of morphologic features in common with the solid tumor tissue he had found elsewhere in the patient's body (6). CTCs arise from the tumor, pass through various intervening structures either actively or passively, and reach the lumen of nearby vasculature. Only approximately 0.0000001% of all tumor cells will reach the bloodstream, according to estimations by Fischer (7). CTCs circulate alongside normal blood cells in the vasculature, where they account for such a tiny percentage of nucleated cells that they are virtually unrecognizable on routine peripheral blood smears and are extremely challenging to detect, even using sophisticated instrumentation. If CTCs are present in a peripheral blood draw, they account for only a fraction of 0.0001% of all nucleated cells (1–10 cells/mL) (8).

Sensitivity of CTC detection in patient blood is highly dependent on the method of detection that is used, as different methods likely detect different subpopulations of these cells. Because the true number of CTCs in patients is unknowable, approximations of sensitivity for each method are generally determined by spiking experiments using a known number of cells from a cancer cell line spiked into healthy donor blood. For the CELLSEARCH® system (Veridex, Warren, NJ), the only FDA-approved CTC detection system, sensitivity of ≥85% has been reported (9). With other methods, like microchips or a platform that considers all nucleated cells, such as the high-definition single-cell analysis (HD-SCA) platform (Epic Sciences, San Diego, CA), sensitivity can be as high as 99.9% (10, 11).

HD-SCA: high-definition single-cell analysis

OS: overall survival

PFS: progression-free survival

ctDNA: circulating tumor DNA

Regarding specificity, healthy people and the limited sets of people with benign disease thus far studied rarely have CTCs (0.3% of samples show ≥2 CTCs/7.5 mL blood sample). Like sensitivity, specificity varies with the method used. Specificity for CTC detection can be as high as 99.7%, as reported using the CELLSEARCH® system in 145 healthy women and 199 women with benign breast disease (9), or even up to 100% with a microchip platform (acquired from 20 healthy subjects, although none with benign disease were studied in this publication) (11).

CTC prevalence in patients who harbor malignant disease differs with carcinoma type and stage. In early breast cancer, CTCs are detected in 18–30% of patients compared with detection rates of ~70% in patients with metastatic disease (12). A study using CELLSEARCH® and blood samples of metastatic patients with colon, breast, rectal, gastric, ovarian, and prostate cancer resulted in 54% of total samples with detected CTCs, including 71% CTC-positive samples for breast cancer, 64% for colon cancer, 33% for gastric cancer, 66% for rectal cancer, 60% for ovarian cancer, and 20% for prostate cancer (13). Another study showed similar results for CTC occurrence in breast cancer, with a detection rate of ~70% in metastatic breast cancer patients in a large cohort (14).

The prognostic potential of CTCs has been discussed since the 1960s, when they were reported to be associated with decreased overall survival (OS) (15). However, other CTC studies between 1955 and 1962 resulted in varying incidences of CTCs in the peripheral blood, and this disparity led to an official statement by the World Health Organization (WHO) in 1963 that CTCs were not yet established as a detection or diagnostic method and needed further studies before clinicians should draw conclusions (16, 17). This and other interesting developments in CTC history over the last 150 years are summarized in **Table 1**.

Today, studies confirm the correlation of CTC presence with decreased OS or decreased progression-free survival (PFS) in many cancer types, including breast, colon, lung, ovarian, and prostate cancers (8, 18). The National Comprehensive Cancer Network even defined a new category of disease for CTC-positive breast cancer patients in its 2015 clinical practice guidelines on breast cancer (19).

CTCs have been demonstrated in various cancer types, and numerous studies show an association of CTC presence with poor prognosis for patients with melanoma (tumors derived from melanocytes) (20) and sarcoma (tumors of mesenchymal origin) (21). However, the best-developed applications and methodologies are those for carcinomas, and we focus on this area in this review.

The concept of a fluid biopsy is not limited to evaluation of circulating cells derived from tumor. Other components of tumor-produced material, or tumor-associated material, are often discussed under this broad umbrella. For example, a lot of exciting work is currently being done in evaluation of isolation of circulating tumor DNA (ctDNA) (22), exosomes, and platelets (23). Platelets accompany and interact with CTCs through direct signaling in the bloodstream (24). Exosomes are membrane-bound vesicles and are part of the signaling process for intercellular molecular communication. Tumor-derived exosomes carry DNA that reflects the genome and mutational status of the tumor and therefore have potential as biomarkers for cancer and metastasis detection

Table 1 Overview of some milestones in CTC research history

Year	Milestone	Reference(s)
1841	Langenbeck performed a microscopic study of cancer patient blood. He reasoned that cancer cells are carried by blood.	163
1869	Ashworth observed cells in the blood that were similar in appearance to those in the primary tumor of a woman with metastatic breast cancer.	6
1889	Paget developed his seed-and-soil theory: Metastasis depends on the interaction between a CTC (seed) and an organ-specific microenvironment (soil).	96
1955	Engell used saponin to lyse red blood cells and produced white blood cell concentrates with CTCs. He found 50% positive samples for advanced carcinoma samples.	164
1959	Seal developed a floatation method with silicone oil (45% of samples contained CTCs).	165
1960	Alexander & Spriggs studied white blood cell concentrates by the dextran sedimentation technique. Slides were prepared and analyzed in bright field. CTCs were compared with other cell types in the blood.	166
1962	The first long-term follow-up study with 99 patients (with colon and breast cancer) found that overall survival correlated with the absence of CTCs.	15
1963	The WHO warned that CTCs are unreliable and that more studies would be needed before "drawing any conclusions" for clinical decision making.	16
1975	A review of CTCs by Salsbury summarized results of CTC research and the role of CTCs in metastasis. The review pointed out varying results of correlations of CTC occurrence with overall survival, drawing the conclusion that CTCs are not a death sentence.	16
1987	CTCs were isolated through buffy coat, and their presence was correlated with metastasis/micrometastasis or disseminated single tumor cells. The method used was Ficoll density gradient plus cytokeratin stain.	167, 168
1993	CTCs were detected using immunobead-PCR: Dynabeads® labeled with Ber-Ep4 (anti-EpCAM).	169
1998	The first trials of DAPI-positive, cytokeratin-positive, CD45-negative cell isolation with flow cytometry occurred.	170
2000	ISET® (isolation by the size of epithelial tumor cells) was developed.	171
2001	Terstappen combined EpCAM detection of CTCs with flow cytometry.	152
2004	The FAST system (a fiber-optic array scanning technology for direct CTC analysis) was developed.	172
2004	CELLSEARCH® was introduced (CTCs "represent a . . . 'real-time' biopsy"), with FDA approval obtained on January 21, 2004.	173, p. 826
2007	A microfluidic device to capture CTCs, the CTC-Chip, was introduced.	11
2010	HD-CTC (direct CTC analysis through digital microscopy scanning) was introduced.	174
2014	Yu et al. and Hodgkinson et al. developed CTC-derived explants to optimize studying the tumorigenicity of CTCs and individual drug sensitivity.	137, 175
2015	Vita-Assay® (functional capture of invasive CTCs by preferential binding of these cells to a mimicked tumor microenvironment) was introduced.	176

or monitoring (25). Additionally, fragments of ctDNA may be released to circulate freely in the bloodstream by apoptotic or necrotic CTCs, primary tumor cells, or cells from a metastatic site (22). However, white blood cells (WBCs) also release DNA and therefore contribute to the bulk of DNA detected in the bloodstream, which makes the specific detection of ctDNA challenging. Ultrasensitive sequencing techniques allow mutation identification with 96% specificity in late-stage disease, but very low concentrations of ctDNA in the total circulating DNA fraction lowered detections in stage I non–small cell lung cancer (NSCLC) to 50% (26). Specificity is also

NSCLC: non–small cell lung cancer

a challenge, but ctDNA is nonetheless another biomarker that is obtainable from a blood sample, providing information about genomic aberrations affecting the efficiency of targeted drugs or revealing the effect of chemotherapeutics on the genome (22).

PHYSICAL PROPERTIES OF CIRCULATING TUMOR CELLS

Diameters of Circulating Tumor Cells

CTCs can derive from many different carcinoma types, and their size varies depending on the tissue of origin or possibly other factors such as changes in protein expression. Small cell lung cancer cells might reach a diameter of only 10 μm, whereas breast cancer CTCs might have a diameter of up to 70 μm, and prostate CTCs can occasionally be larger than 100 μm (27). The surrounding WBCs—erythrocytes (~8-μm diameter), granulocytes (~15-μm diameter), lymphocytes (6–18-μm diameter), and monocytes (12–20-μm diameter)—are generally smaller (28).

Deformability of Circulating Tumor Cells

Deformability (the ability to change shape under the application of stress) can be measured using a microfluidic optical stretcher, which is a two-beam laser trap that deforms single cells in suspension by optically induced surface forces. A mixture of cell line CTCs can be sorted according to whether they derive from benign (MCF-10), nonmotile/nonmetastatic (MCF-7), or highly metastatic (MDAMB-231) cells on the basis of varying optical deformability, with cells derived from highly metastatic tumors showing increased deformability (29). The difference in deformability results from various factors like extracellular matrix (ECM) signaling, chemotherapy, or DNA rearrangements that may cause a shift in protein expression. This shift transforms the structure of the cytoskeleton in cancer cells, causing their shape, motility, and deformability to differ as well (30). The nucleus of a cancer cell is often enlarged, which causes the cell to be more rigid than a WBC. In 2016 Park et al. (31) enriched CTCs from the blood of prostate cancer patients on the basis of their lower deformability relative to blood cells, with a 25-times-higher yield relative to enrichment by surface epithelial marker positivity using CELLSEARCH®. In cancer cell lines, in contrast, higher deformability has been linked to higher metastatic potential of the tumors from which the cell lines were derived (29, 32), and in studies with patient-derived CTCs, the comparison of benign epithelial cells with CTCs demonstrated higher deformability of CTCs (33). Therefore, the literature seems to conclude that CTCs are less deformable than WBCs but show increasing deformability as they increase in metastatic potential (34).

Polarity and Electrical Charge of Circulating Tumor Cells

As a result of changes in shape, nuclear size, and cytoplasm as well as protein expression changes during mobilization and/or activation of survival mechanisms in circulation, the ratio of suspensoid polarizable particles (e.g., proteins, nucleic acids, and peptides) to their solvent is different in CTCs than in WBCs and benign epithelial cells. This altered ratio becomes an important and very useful dielectric property (35). One method to exploit electrical properties is dielectrophoresis (DEP), in which a nonuniform electrical field separates cells by either pulling them away or moving them toward the electromagnetic field, depending on the cells' own electrical charge (35). Gascoyne & Shim (36) demonstrated a lack of overlap in dielectric properties of cancer cells from cell lines versus blood cells so that DEP can be used for CTC enrichment. The isolated cells were transferred to cell culture medium and demonstrated viability. Proof of concept in multiple cancers was shown in a study with patient samples from NSCLC, breast cancer, and prostate cancer; dielectric properties were successfully used to enrich CTCs (37).

Density of Circulating Tumor Cells

Other normal constituents of whole blood provide a convenient internal control against which to measure various properties of CTCs. A predominant cell type in whole blood is red blood cells (RBCs). Their cytoplasm contains hemoglobin, a protein bound to iron atoms, which causes RBCs to have higher density (weight per volume) than WBCs or CTCs (38). In 2012 Phillips et al. (39) compared densities (cellular dry mass density) of CTCs, RBCs, and leukocytes in an ovarian cancer patient. They reported not only 3.5–4.5-times lower densities of leukocytes and CTCs compared with RBC density but also a density overlap within the first two types of cells. Thus, CTCs would reside in the buffy coat in the clinical lab.

BIOLOGICAL PROPERTIES OF CIRCULATING TUMOR CELLS

Epithelial Adhesion Markers and Cytokeratins

In carcinoma patients, the most commonly used cell surface markers for the enrichment of CTCs from a blood sample is the epithelium-specific cell adhesion molecule (CAM) EpCAM, which contributes to cohesion within an intact epithelial structure. Cytokeratin (CK) is then often used to further confirm the epithelial nature of enriched or captured cells.

Tumor-Specific Markers

CTCs generally retain the protein expression profile of the tissue from which they are derived. This proteomic profile not only is epithelium specific but also, unless the tumor is extensively dedifferentiated, contains an organ-specific proteomic signature. In tissue biopsies this characteristic is exploited by using stains such as differential CKs (e.g., CK7, CK20) or organ system–specific stains like prostate-specific antigen (PSA), thyroid transcription factor 1 (TTF-1), and the homeobox protein CDX-2. Clinically actionable markers, such as receptor proteins that are used in tissue biopsies to guide therapy, can be found in CTCs as well. These markers and their expression levels are used to validate detection methods for CTCs, and their potential to add prognostic information over and above current standard tumor markers from serum or tissue is being investigated. Studies are currently evaluating the comparability of marker expression in CTCs with the primary tumor or metastatic lesion (40). For example, the expression of receptors such as HER2 in metastatic breast cancer (41) and EGFR in colorectal cancer patients (42) can exhibit disparities between CTCs and the primary tumor. Such disparities may reflect an evolution of disease or may represent the true heterogeneity of a tumor that underwent only limited sampling. Today's knowledge about proteomic profiles of CTCs is insufficient to guide therapeutic decisions, and thus only a few early trials are testing therapy alteration based on marker levels in CTCs, such as HER2 expression in metastatic breast cancer CTCs (43–45).

MORPHOLOGIC HETEROGENEITY OF CIRCULATING TUMOR CELLS

Because different methodologies capture varying subsets of CTCs, it is difficult to fully characterize the breadth of the morphologic heterogeneity of these cells. Techniques capable of detecting the widest spectrum of possible CTCs have the most potential to capture more of this morphologic heterogeneity (46, 47). With the HD-SCA platform, for example, all nucleated cells from a blood sample are preserved for analysis. "No cell gets left behind," and from each candidate cell, images are generated from three or four fluorescent channels, each devoted to a separate protein.

Heterogeneity of Circulating Tumor Cells and Circulating Tumor Cell Candidates Within Single Samples and Within Single Cancer Types

CTCs are not uniform, even within a single sample or within a single cancer type. Under investigation are not only traditional CK-positive, CD45-negative cells with a strong DAPI signal (HD-CTCs), but also other, less-understood cells that are captured by detection systems and that morphologically or immunophenotypically differ from cells of the hematopoietic lineage (see **Figure 1**). Conventional HD-CTCs show a wide variety of shapes and sizes, with a distinct cytoplasmic CK signal surrounding a nucleus that is larger than the surrounding WBC nuclei. Other categories of cells captured and annotated with the HD-SCA platform, and described as well by other investigators, are CTCs with low or no CK signal (CTC-LowCK), CTCs with a nucleus that does not differ in size from WBC nuclei (CTC-Small), and CTCs that show apoptotic characteristics such as a disrupted nucleus and/or CK stain and that are theorized to be a source of tumor DNA [CTC–cell free DNA (cfDNA) producing] (48). CTC-LowCKs show an enlarged, irregularly shaped nucleus. The loss of CK signal in these cells may be due to a transition to a mesenchymal phenotype, which would imply that these cells, generally not identified with EpCAM-based methods, may be more invasive and thus even more important for the prediction of disease progression than their epithelium-like counterparts (49). Preliminary data in breast cancer have also shown that these cells indeed frequently have an aberrant copy number variation (CNV) profile, and further investigations using single-cell next-generation sequencing (NGS) will give more insight into this cell category. Alternative possibilities include stripped cells with absent cytoplasm resulting from sample preparation effects. Such cells may also represent an entirely different cell type, for example, immature megakaryocytes. The next category of cells that are of potential interest are those that appear apoptotic (labeled in **Figure 1** as CTC-cfDNA producing). Tracking of these cells potentially allows for evaluation of the extent of dying CTCs in the blood, with potential treatment response implications. Another category of small but CK-positive and CD45-negative cells is confounding to many groups, particularly because in tissue biopsies of cancer, tumor cells are virtually always larger than any surrounding benign cells.

NGS: next-generation sequencing

Heterogeneity of Circulating Tumor Cells in Cancer Types

Finally, comparison of intertumor morphologic features of CTCs may be studied as well. Park et al. (50) observed, in prostate cancer patients, small CTCs only ~8 μm in diameter (similar in size to leukocytes) but with an elongated shape (50). These findings concur with those from another study that demonstrated that CTCs of prostate and colorectal cancer are smaller in size than breast cancer cells (51). **Figure 2** shows some differences in CTC morphology in four cancer types. Thorough and extensive comparisons of the morphologic features of CTCs from various tumor types, with attention to the grade(s) and specific histologic features of the tumors from which they have arisen, are lacking, but such comparisons may emerge as techniques for isolation and visualization of these rare cells become more available.

KINETICS OF CIRCULATING TUMOR CELLS

There are many theories about the origination of CTCs. What mechanisms are in play as these anchorage-dependent tumor cells leave the tumor to face harsh hemodynamic stress, immune response, a CTC-hostile microenvironment, and probable death (52)? How do they get into the bloodstream, how do they get out, and how long do they circulate? Experimental evidence has been collected over the years, and although mysteries remain, some progress can be reported (**Figure 3**).

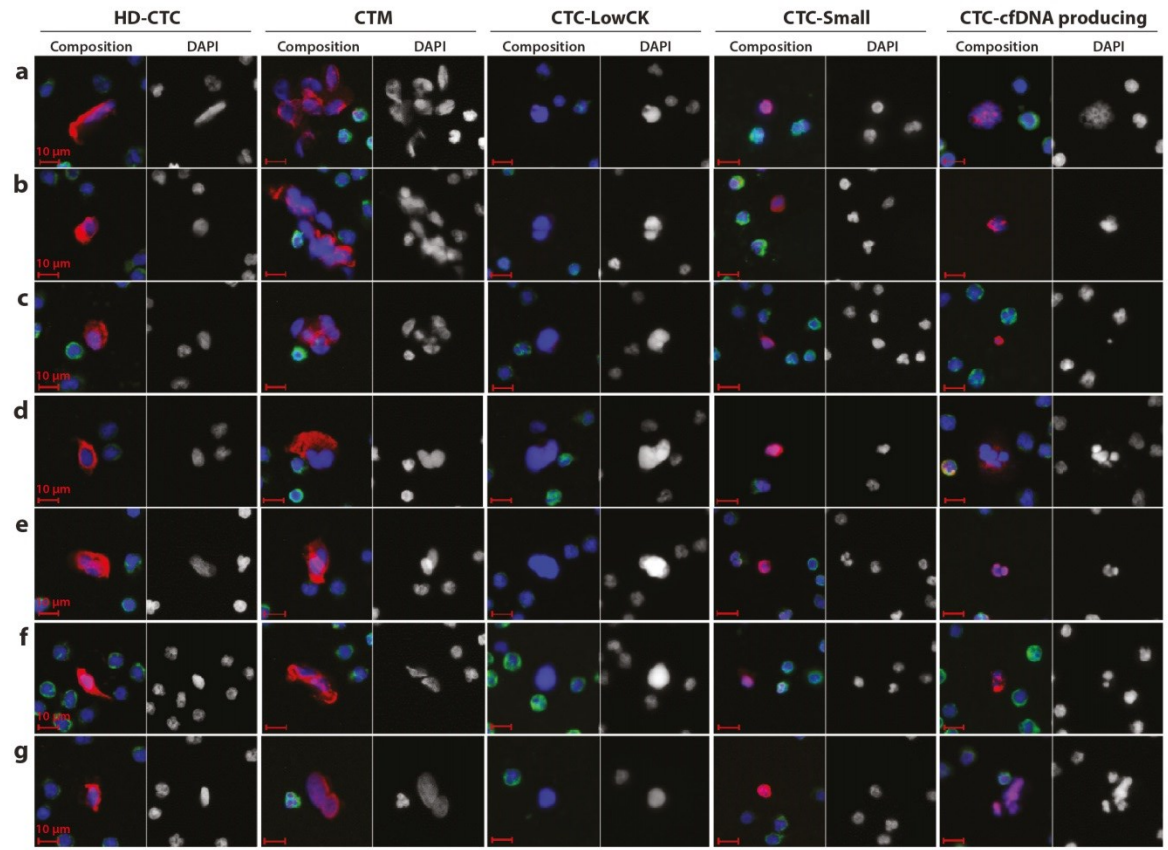

Figure 1

Candidate cells of circulating tumor cell (CTC) detection by the high-definition single-cell analysis (HD-SCA) approach of Kuhn and colleagues (J.-A. Thiele, P. Pitule, P. Ostašov, V. Liška, M. Králíčková, et al., unpublished data). The images represent the pleomorphic population of candidate cells found in the blood (before surgery) of metastatic colorectal cancer patients; these cells are stained with DAPI (*blue*), CK-pan mix (*red*), and CD45 (*green*). The composite image and the DAPI image are displayed for each cell type. Cell type abbreviations (*from left to right*): CTCs detected by the HD-SCA platform (HD-CTCs), CK-positive and CD45-negative cells with a nucleus distinct from WBC nuclei; circulating tumor microemboli (CTM), HD-CTC clusters (at least two or more HD-CTCs); CTC-LowCK, cells with a nuclear shape different from that of WBCs but CK negative and CD45 negative; CTC-Small, CK-positive and CD45-negative cells with a small (WBC-like) nucleus; CTC-cfDNA producing, CTCs undergoing apoptosis with irregular nuclear or cytoplasmic condensation and a possible source of circulating tumor DNA.

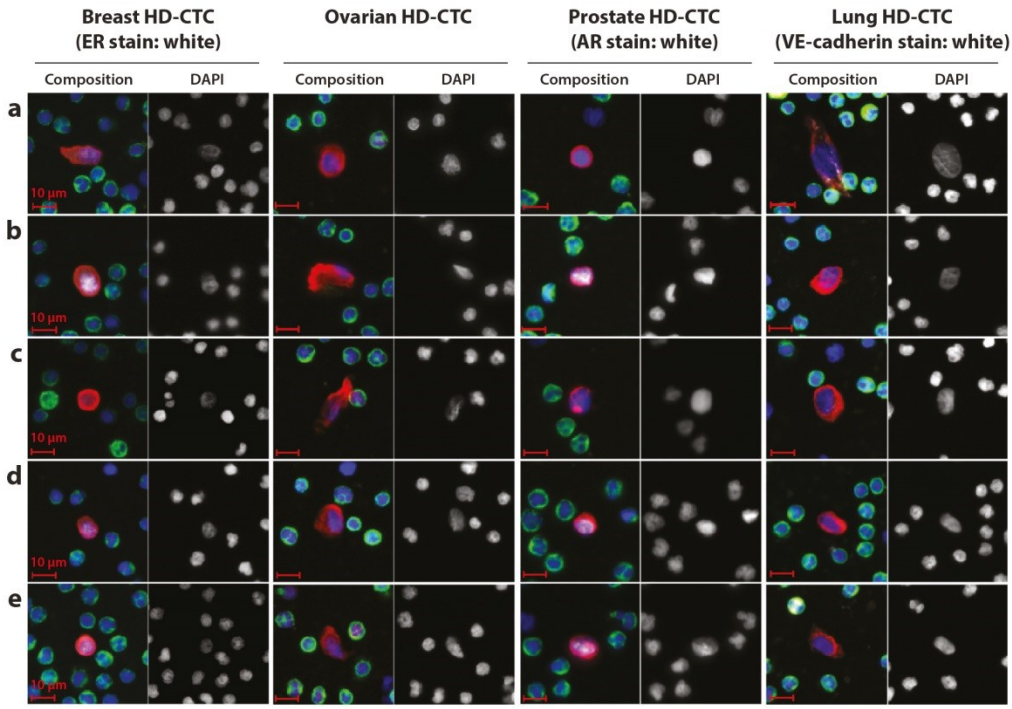

Figure 2

Circulating tumor cells (CTCs) detected by the HD-SCA platform (HD)-CTCs in four different cancer types (C. Ruiz Velasco, L. Welter, N.A. Carlsson, P. Malihi, P. Kuhn, et al., unpublished data). Although this figure does not thoroughly classify CTC morphology differences by cancer type, these images demonstrate some subtle but recognizable differences. In breast HD-CTCs, the endoplasmic reticulum shows a very distinct pattern (*b*), and these CTCs are more round in shape than the HD-CTCs of other cancer types. For ovarian HD-CTCs, variation in shape is high. Prostate HD-CTCs are round in shape (similar to breast HD-CTC shape) and are often not much bigger than white blood cells (results are concordant with Reference 50). Lung HD-CTCs show more elliptical shapes and varying sizes compared with the other three HD-CTC types shown here.

Passive Shedding of Circulating Tumor Cells

Many research groups have investigated the existence of passive tumor cell shedding into the bloodstream (53, 54). Even with complex experimental setups, distinguishing between passive shedding and active intravasation is difficult. Therefore, the theory of passive shedding as a frequent mechanism producing CTCs is supported only by circumstantial evidence. Pathologists routinely observe microscopic invasion of small blood vessels on histologic slides. Such invasion appears to occur via tongues or clusters of contiguous tumor cells growing directly into blood vessels. This commonly observed phenomenon suffers both from being a rather gross observation in terms of individual cell behavior and from being a moment frozen in time at the point of tissue fixation, and abstracting more detailed mechanistic information is difficult. Abundant mosaic blood vessels (endothelial and tumor cells constituting the luminal surface) in various tumors indicate an erosion-type mechanism. A study of colorectal carcinoma showed that 13.4% of the 367 vessels analyzed were mosaic (55).

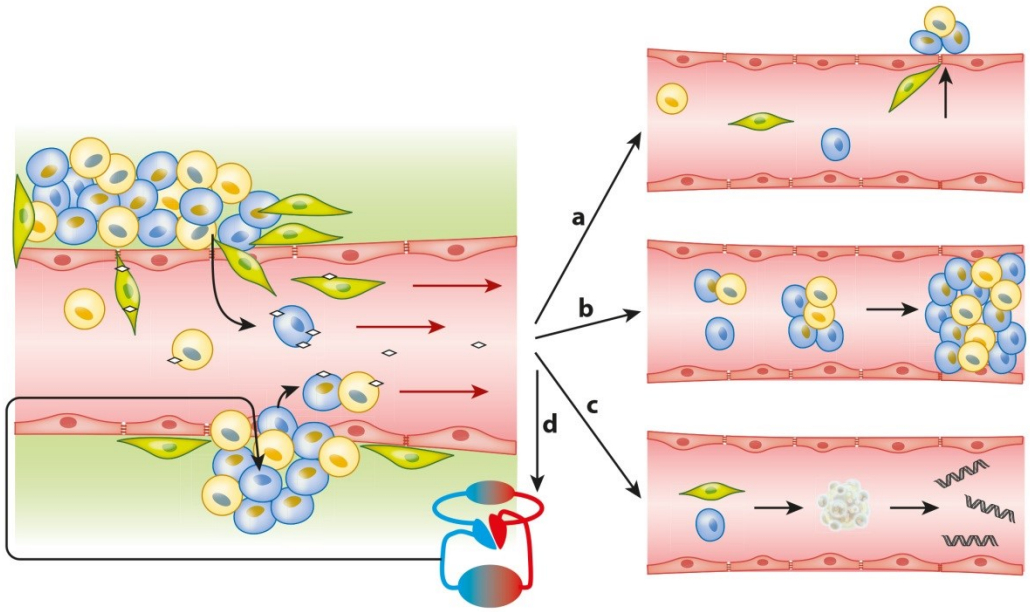

Figure 3

Proposed circulating tumor cell (CTC) dissemination into the vasculature and subsequent prospects. (*Left*) Epithelial heterogeneous tumors shedding tumor cells into the bloodstream. Such shedding can be passive (*bottom*) as parts of the tumor (fragile vessels) break off, thereby shedding epithelial CTCs (*blue* and *yellow*) into the blood, sometimes even as circulating tumor microemboli (CTM). Alternatively, shedding may be active (*top*) as cells undergo epithelial-to-mesenchymal transition (EMT) (*green*, EMT-CTCs), which may happen partially or completely and is often induced by platelet interaction (*little white diamonds*) (24). As these EMT-CTCs actively penetrate into the blood vessel, some epithelial cells may passively follow (73), driven by the bloodstream-induced low pressure. Once within the vasculature, a CTC has multiple fate options. (*Right*) (*a*) Cells that underwent partial or full EMT may reverse this process and start mesenchymal-to-epithelial transition as part of metastasis formation (73). However only 0.01% of CTCs are likely to form a metastasis at a distant site (92). (*b*) CTM and CTCs in the vasculature may get stuck in blood vessels and are suspected to have a role in causing venous thromboembolism (83). Additionally, CTM may initiate metastatic growth after lodging in a distal capillary (74, 87). (*c*) CTCs (epithelial or mesenchymal) may undergo apoptosis, which releases circulating tumor–derived DNA into the bloodstream. (*d*) After survival in the vasculature, CTCs may self-seed. This process is often observed and is supported by the CTC-friendly microenvironment of the primary site (95, 97).

Active Migration into Blood Vessels

Possible more active mechanisms of vascular intravasation include macrophage interactions with tumor cells; such interactions may induce tumor cell movement along collagen fibers toward blood vessels. Tumor cells and macrophages collaborate in a paracrine loop in which the cancer cells express the epidermal growth factor receptor (EGFR) and secrete colony-stimulating factor 1 (CSF-1), which attracts macrophages, whereas macrophages secrete epidermal growth factor (EGF), which binds to the corresponding receptor of the cancer cell. The collagen fibers along which the macrophages and the attracted tumor cells are moving can be envisioned as a spider web, with the center of this web being the blood vessel. Growth factor and nutrient gradients

are directing the tumor cells toward the vasculature (56). This movement of tumor cells along collagen fibers is often described as occurring through a crawler mechanism (57).

The invasive movement of epithelial cancer cells away from their collective and through stromal tissue and vascular walls is an area of extensive research. One of the main theories regarding the mechanisms for such invasion is epithelial-to-mesenchymal transition (EMT). EMT is frequently theorized to help explain active intravasation mechanisms of tumor cells into the vasculature. In this process, epithelial tumor cells lose their characteristic cell-cell contacts and apical-basal polarity to gain elongated morphology with increased mobility and invasive properties (58). Activation of this process, which leads to a more invasive cell type, is thought to be induced by the microenvironment of the tumor (59). EMT-triggering signals include transforming growth factor β (TGF-β) secretion by platelets (24), integrin-macrophage interactions through EGF supply (as mentioned above) (56), secretion of proinflammatory cytokines by fibroblasts (60), and hypoxia (the tumor outgrowing its blood supply) (61), in addition to others (62, 63).

As a result of the EMT-triggering signals, transcription factors like snail family transcriptional repressor 1 (SNAIL), snail family transcriptional repressor 2 (SLUG), and forkhead transcription factor 2 (FOXC2) are activated (58). A series of signaling networks are induced and cause, for example, the loss of the CAM E-cadherin. E-cadherin downregulation triggers the so-called cadherin switch and enhances expression of N-cadherin, a promoter of motility and invasion (64). CK expression is also reduced, whereas expression of vimentin and many more key regulators of EMT—like the helix-loop-helix protein Twist, E-box-binding repressors Zeb1 and Zeb2, β-catenin, Rac1, Akt2, and the cytoskeletal regulator integrin—increases (65–67). All these factors finally cause a redesign of the cytoskeleton, mainly by a switch from a CK-rich filament network to one composed mainly of vimentin (68). Adherent junctions are destabilized, polarity is lost, and filo- and lamellipodia for migration form. The cell finally reaches the mesenchymal, mobile, and protected phenotype, which allows for invasion through stroma and possibly intravasation and survival in the bloodstream (66, 68). All these factors are being investigated as markers for determining the cell state and invasive character of isolated CTCs.

Diagnostic pathologists recognize tumor types that have characteristics of EMT. Certain tumors are known to have lost some of their cohesive epithelial nature. This loss can be immunohistochemically demonstrated; one example is the loss of E-cadherin staining in lobular breast cancer and certain diffusely infiltrating gastric cancers. Within an individual tumor, however, histologic identification of specific cells or areas that may be undergoing EMT at the invading front of a tumor is challenging from a pathologic standpoint and has not historically been pursued. These cells usually do not display a typical spindle-shape mesenchymal phenotype and for that reason are rarely observed in the clinical histopathologic environment (69). However, in 2014 Ueno et al. (70) developed an approach to visibly detect and evaluate EMT potential in colorectal cancer histologically. They developed the Histology^{EMT} system by using poorly differentiated clusters, desmoplastic reaction, and stroma maturity to analyze EMT at the leading edge of the primary tumor and showed that this method has significantly higher prognostic value than the current standard of TNM staging for disease-free survival (70). Recent papers evaluated the leading invasive edge of various cancers by immunostain panels and describe an immunohistochemically recognizable diminution in the staining intensity of some surface epithelial markers with increased vimentin or other mesenchymal marker positivity in cells at the leading invasive edge relative to those within the central tumor mass. In fact, the plasticity of cells at the leading invasive edge has been proposed to be an inherent quality that may be a better indicator of invasive and metastatic potential than overall tumor differentiation (71, 72).

EMT may also be secondarily involved in the passive shedding of tumor cells. Tsuji et al. (73) found evidence for a synergy between cells undergoing EMT and cancer cells that do not

(non-EMT cells). The cells undergoing EMT degraded the surrounding matrix, enabling invasion and intravasation of non-EMT cells.

CTM: circulating tumor microemboli

VTE: venous thromboembolism

Even though EMT has been theorized as a leading mechanism of metastasis, this theory has also been debated (73) and recently even challenged. Aceto et al. (74) revealed that EMT is not necessarily responsible for the mobility of tumor cells. In addition, recent mouse model experiments in which EMT regulators were downregulated showed that EMT was dispensable for metastasis but contributed immensely to chemotherapy resistance (75, 76).

Survival of Circulating Tumor Cells in the Vasculature

After intravasation, CTCs are confronted with multiple stress factors. The sheer force generated by the bloodstream is quite different from the forces exerted on cells in their tissue of origin. In addition, epithelial cells that lose their cell-cell and cell-matrix interactions usually undergo anoikis (77). CTCs may also undergo anoikis and release cfDNA (22). Studies of kinetics suggest a short survival time in the bloodstream. A mouse model experiment with renal cell carcinoma cell lines demonstrated total cell numbers shed (as directly measured from the renal vein of the kidney) to be as high as 6 million cells/(day·g), with approximately 89% of these cells being nonviable immediately after being shed and a predominant subset of the nonviable cells being apoptotic (78). The viable cells usually have a short lifetime. Meng et al. (79) measured, in breast cancer patients, a CTC half-life of 2.4 h, with a maximum survival of 24 h after primary tumor resection. Rossi et al. (80) recently detected and distinguished between spontaneous and drug-induced apoptosis of CTCs in breast, renal, and colorectal cancer patients. Using M30, a biomarker for apoptosis in epithelial cells, Rossi et al. detected a fraction of 50–80% M30-positive CTCs undergoing spontaneous apoptosis.

Circulating Tumor Cell Aggregates and Their Possible Contribution to Venous Thromboembolism

CTCs are not found just as single cells within the vasculature. They can often be detected as aggregates, so-called circulating tumor microemboli (CTM), which are composed of two or more cancer cells (81). Studies on mouse models suggest that, rather than CTCs forming clusters after intravasation, the aggregates of cells enter the vasculature as groups after dissociating from a solid tumor mass (74). This process has also been associated with overexpression of vascular epidermal growth factor A (VEGF-A) by tumor cells; VEGF-A may induce the release of CTM into the vasculature (82). CTM are theorized to play an important role in both tumor metastasis and venous thromboembolism (VTE), which are the two main causes of mortality in cancer patients (83). Cho et al. (81) detected CTM in 54% of breast cancer patients, in 52.5% of NSCLC patients, and in 73% of prostate cancer patients (all stage IV patients). Frequently, both CTM and CTCs were physically associated with other cell types such as fibroblasts (84), leukocytes (85), endothelial cells (86), and platelets (24). Experimental evidence has supported the view that CTM may be particularly ominous; as early as the 1970s, CTM were shown to have a higher metastatic potential in mouse model experiments comparing tumor growth from injection of single cells with tumor growth from injection of clusters (87), an observation recently confirmed by Hou et al. (88).

During circulation through blood vessels, CTCs and most notably CTM may contribute to VTE (83), but the mechanisms are poorly understood. The expression of the coagulation cascade member tissue factor (TF) by CTCs may play a role. TF is a trigger of coagulation and can be regulated by oncogenes, tumor suppressors, or growth factors in cancers (89). In 2014 Phillips et al. (90) demonstrated, through a dynamic simulation of CTCs and CTM under flow within the vasculature, that CTM may cause a regional elevation of thrombin levels that is enabled by

reduced flow speed caused by vessel junctures. CTCs and CTM have thus been suggested as mobile sources of thrombin, but their detailed role in VTE remains ill defined (91). Given cell counts of up to 50 within one CTM, another theory for VTE could be based on a pure size issue, with the cell mass clogging a small blood vessel, as displayed in **Figure 3*b*** (74).

Establishment of Metastasis

The last step of metastasis is the actual occupation of distant areas of preexisting normal tissue, and this development requires exit from the bloodstream and the establishment of a stable proliferative state in a new location. This process is highly complex and inefficient. In animal models only 0.01% of all tumor cells entering the bloodstream were able to extravasate into tissue and form a metastasis (92). Other potential outcomes are anoikis, destruction by immune cells, and transition to a dormant nonproliferating state (93). Whether extravasation can take place is dependent on the environment that CTCs find during their travel and on the inherent metastatic potential of the CTCs. If a metastable CTC arrives by arrest in a small capillary with subsequent ischemic capillary wall destruction, by extravasation through the capillary endothelium, or by extravasation across the wall of a larger vessel, it may encounter a favorable microenvironment known as the metastatic niche. Some of the factors defining the metastatic niche are stromal cells and ECM proteins that support survival and self-renewal (52). Duda and colleagues (94) demonstrated in mouse models that CTCs may even bring their own "soil" from the primary site. They observed CTCs that were embedded, during circulation, in stromal components, like fibroblasts, macrophages, and endothelial cells, which provided the CTCs with survival and growth advantages. Alternatively, favorable environments may already be present within an existing tumor or metastatic site that a CTC might encounter (95). Therefore, Paget's seed-and-soil theory (96) cannot be seen as an unidirectional process, but rather more as an encounter between a circulating cell and a suitable environment (97).

Finally, because well-established metastases almost always strongly resemble the primaries from which they arose, EMT as a mechanism of circulatory spread must be an evanescent state and requires a reversal step to explain the fully expressed epithelial features recognized by pathologists in established metastases. The cells somehow have to achieve a recognizably epithelial state again. The theory that CTCs undergo a reverse process, the mesenchymal-to-epithelial transition (MET), has been proposed (58). In 2016 this potential reverse transition was further examined, and adenosine 3′,5′-monophosphate (cAMP) and its main effector, protein kinase A (PKA), were proposed to play key roles (98). MET involves activation of extracellular proteases that dissolve membrane proteins, as well as reactivation of CAMs, integrins, E-cadherin, and β-catenin (99), as portrayed in **Figure 3*a***.

OTHER EMERGING CONCEPTS

Spreaders and Sponges

The journey of a CTC may not be a one-way trip from a primary tumor to metastatic niche (100, 101). Newton and colleagues (102) developed a mathematical (Markov chain) model, attempting to shed light on the complex circulation pathways of CTCs. They described a pattern of metastasis best explained by three pathways: self-seeding of the primary tumor, reseeding of the primary tumor from a metastatic site, and self-seeding of metastatic tumors. Using a large autopsy data set from lung cancer patients, they were able to classify metastatic sites into two distinct categories. If a metastatic site appeared to absorb more metastatic potential than it gave rise to, then it was classified as a sponge. If the site was observed to spread more to other sites, then it was classified

as a spreader. It was determined from the model that an important metastatic site, and the most important sponge, was the group of regional lymph nodes. Although the model is still being developed, it is able to illustrate how cancer progression is a multidirectional process (102). In the future, these types of models could potentially be utilized for prediction of metastatic sites that will inform clinicians when and how to adjust treatment.

Dormancy

How can patients go for many years with no evidence of disease and yet relapse eventually with the same cancer? Moreover, patients can have evidence of CTCs even years after surgery without disease recurrence (79). These observations suggest that there is a condition of CTCs in which extravasation, or growth of extravasated tumor cells into metastases, is temporarily stopped. This state is known as dormancy and is proposed as a mechanism often utilized by cells to adapt to new microenvironments (103). Dormancy is defined as a nondividing phenotype. Ki-67, a nuclear protein for cell proliferation, is absent or downregulated in CTCs that are considered dormant (104). Studies have shown that, even up to 22 years after mastectomy, dormant CTCs can be found in patient blood (79). The time from dormancy to metastatic progression has been referred to as the latent niche (105). Whether dormancy (Ki-67-negative CTCs) could be related to future relapse has not been fully established. However, authors of a study of breast cancer patients undergoing adjuvant chemotherapy observed that Ki-67-negative CTCs may be resistant to therapy (106). A better understanding of this process and identification of dormant cells would enable the targeting of these cells, which may represent a type of minimal residual disease.

TECHNIQUES: METHODS FOR CAPTURE AND CHARACTERIZATION OF CIRCULATING TUMOR CELLS

Today's CTC technologies are pushing past mere identification and enumeration and are attempting to better understand the molecular features of these rare cells in a two-stage process. The first stage is a detection or capture stage, often utilizing some type of enrichment component that increases CTC concentration and/or depletes surrounding normal blood cells. In the second stage, either a CTC-enriched population of cells or a pure population of individually retrieved CTCs is further characterized by various molecular techniques.

To date, detection methods have been developed to target specific biological and physiological properties of CTCs. Initial specificity, efficiency, and purity measures are usually obtained by spiking cancer cell line cells into healthy donor blood followed by further validation with clinical specimens. Multiple methods are available and have been discussed at length elsewhere (107–109). Here, we provide a brief overview of different approaches for CTC detection (see **Figure 4**), including established methods as well as novel techniques that are under development.

Positive Detection Strategies

CTC enrichment can be performed by utilizing EpCAM as a positive detection strategy. All EpCAM-based technologies focus on this cell adhesion protein expressed on the surfaces of epithelium-derived CTCs. The detection step is based on antibody-labeled beads, columns, microposts, or other devices to isolate CTCs.

A commonly recognized method for CTC detection is CELLSEARCH®. An automated system for enumeration of CTCs, it uses immunomagnetic capturing of cells that express EpCAM on the cell surface; in this system, a CTC is defined as a nucleated DAPI-positive, CK-positive, and CD45-negative cell (110). CELLSEARCH® is the only FDA-approved CTC detection method.

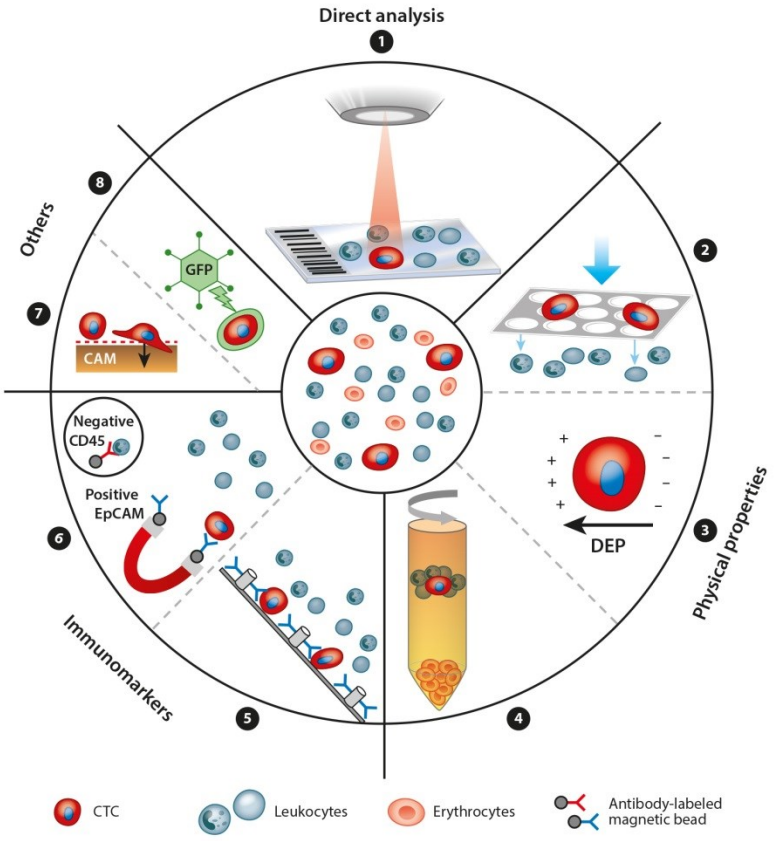

Figure 4

The principles of circulating tumor cell (CTC) enrichment. CTCs can be detected in many ways. A nonenrichment approach involves the direct analysis, by microscopy, of all peripheral blood mononuclear cells of a patient sample (❶). The enrichment approaches are divided into label dependent (addressing protein expression) and non–label dependent (enrichment by physical properties of CTCs) approaches. Positive or negative enrichment by physical properties (❷❸❹) is possible, for example, using filtration by size (e.g., ISET®) (❷), using dielectrophoretic properties (DEPs) (❸), or using centrifugation through a Ficoll density gradient (❹). Enrichment through immunomarker positivity (e.g., for epithelial markers like EpCAM or mesenchymal markers like N-cadherin) is a common and well developed method (❺❻). Label-dependent techniques have been optimized by many microdevices like the CTC-Chip by using antibody-labeled microposts (❺). Alternatively, negative enrichment may be performed by depletion of leukocytes using antibodies against CD45 (❻). Some other approaches on the market use specific CTC functions like protein secretion (EPISPOT) or invasive behavior on a cell adhesion matrix (CAM) (❼). DNA targets can be detected within CTCs through aptamer technology or through oncolytic green fluorescent protein (GFP)–containing, telomerase-specific adenoviruses (❽).

Enumeration of CTCs has been proven to be a prognostic marker for metastatic breast, colorectal, and prostate cancers (8). CELLSEARCH® has opened the door for CTCs as potentially valuable clinical biomarkers. However, the mere enumeration of CTCs has proven insufficiently informative to engender widespread clinical adoption.

Other approaches include microfluidic platforms like the CTC-Chip. CTCs interact with EpCAM-coated microposts under laminar flow conditions (11). The process involves a one-step setup and is very gentle with regard to cell handling. Over time, the CTC-Chip developed into the Herringbone Chip after undergoing multiple improvements (111). The addition of a microvortex within the device has optimized antibody affinity. Moreover, changing the shape from sharp grooves to wavy corners has enhanced the EpCAM-coated surface. These adjustments have boosted the purity of CTCs (meaning an absence of nonspecifically captured WBCs) from 25.7% to 39% while retaining an efficiency of 85% cell recovery (112). These microfluidic chips are a cheap and fast method for CTC capture, but as long as they depend on EpCAM detection, they face the same limitations as CELLSEARCH®.

Following the development of coated microposts, nanostructured substrates were also developed. Silicon nanopillars coated with EpCAM antibodies are combined with micromixing abilities. This technique achieves approximately 95% efficiency but processes only 1 mL/h of blood. This system can detect CTCs in 20 of 26 patients compared with detection in 8 of 26 patients with CELLSEARCH® (113).

Another innovation is the MagSweeper® (Illumina, San Diego, CA) (114), in which blood samples are diluted and prelabeled with EpCAM-coated magnetic particles that can subsequently be captured by a magnetic rod sweeping through the sample. The magnetic rod then switches to a washing well, and in a buffer solution an external magnetic field causes the release of labeled cells, which remain viable and unaffected and are therefore transferable to cell culture (115).

The CellCollector® (GILUPI GmbH, Potsdam, Germany) is an inventive device that captures CTCs on the basis of EpCAM surface expression in vivo. Antibodies are coated on a gold-plated nanoguidewire that is inserted into the patient's cubital vein for 30 min through a venous cannula (116). Thereafter, adherent CTCs can be immunocytochemically stained and evaluated. Treatment response in lung cancer may be associated with decreasing CTC counts, and mutational analysis of captured CTCs was possible (117). This method overcomes blood volume limitations but is a rather unpleasant and more invasive procedure for the patient than a blood draw.

All the methods mentioned above are limited by their reliance on enrichment for EpCAM-positive cells, and thus only the EpCAM-positive subpopulation of CTCs are detected through these methods. CTCs that have undergone EMT may downregulate their epithelial markers like EpCAM and E-cadherin, and these cells would thus be invisible to these methodologies (118).

A somewhat broader capturing approach utilizing immunoaffinity is offered by AdnaGen (QIAGEN, Hilden, Germany). With prelabeled magnetic beads and its combination-of-combination principle, the AdnaTest® offers a variety of capturing antibodies (EpCAM, HER2, MUC1). However, even if the limitation might not be as stringent as EpCAM-only-based techniques, it is still a constraint on a defined subpopulation of CTCs that is being captured. An additional drawback of the AdnaTest® is the lysis of captured CTCs, which allows for bulk PCR analysis of cancer-specific biomarkers but negates the possibility of analyzing single cells by methods such as single-cell NGS.

Detection Based on Differential Physical Properties

Alternative methods include density gradient centrifugation like OncoQuick® (centrifugation only; Greiner Bio-One GmbH, Frickenhausen, Germany) (119), filtration systems isolating

CTCs by size [the most famous example is ISET® (isolation by size of epithelial tumor cells, RARECELLS US Inc., Austin, TX)] (120), and microdevices using size and deformability like the ClearCell® FX (Clearbridge BioMedics, Singapore), which shows approximately 80% efficiency (121). Alternatively, CTCs can be enriched by electrical charge, wherein electrodes create spherical dielectrophoretic cages to detect rare cells and sort them by morphologic parameters.

A label-free method using deformability in combination with microfluidic ratchets was recently used to capture CTCs with a 25-times-greater yield than that of CELLSEARCH®. In addition, the captured cells are viable and obtainable for downstream analysis (31).

Another platform consists of a spiral microchannel using Dean drag forces to separate larger cells (CTCs) by strong inertial lift forces. Dean flow fractionation is able to recover 85% of CTCs in spiked blood samples and isolates CTCs in a completely antibody-independent manner at a rate of 3 mL/h. However, only a low purity of ~10% could be achieved. Validation with clinical samples of metastatic lung cancer patients showed a 100% detection rate (122).

Negative Detection Strategies: Leukocyte Depletion

The biantibody leukocyte depletion kit EasySep® (STEMCELL Technologies, Vancouver, Canada) allows CTC enrichment through leukocyte depletion by using CD45-labeled magnetic beads (123). Leukocyte depletion methods have been reported to result in lower purities than purities obtained by using positive CTC selection (109) and are therefore often combined with other enrichment techniques. However, after depletion methods of enrichment, CTCs are viable and can be used for further experiments (124).

Combined Methods

The approaches described above often do not alone obtain the desired purity or specificity; therefore, a combination of methods can be utilized for detection of CTCs. Negative selection by CD45 sorting can, for example, be combined with a functional-type assay (107). Such a technology, EPISPOT, detects protein secretion of viable CTCs through labeled antibodies on the bottom of the culture dish (125). Through this method, researchers were able to classify metastatic breast cancer patients into high- and low-risk groups. In combination with the CELLSEARCH® method, reported data were a strong predictor for OS (126). Another combinational platform is maintrac (Simfo GmbH, Bayreuth, Germany), which consists of RBC lysis, centrifugation, and detection through EpCAM-labeled antibodies (127). Another example is CanPatrol® (SurExam Biotech Ltd., Guangzhou, China), which involves CTC enrichment by filtration followed by RNA in situ hybridization. For example, by using EMT-related markers like Twist and vimentin, recovery rates of 80% were reported (128).

A combination strategy consisting of microfiltration and a telomerase-selective adenovirus takes advantage of the enhanced telomerase activity of tumor cells. Experiments were executed without filtration, but the viruses were not specific for cancer cells, and therefore filtration by size was included as an enrichment step. With this strategy, sensitivity was between 75% and 93%, but false-positive cells were detected as well (129). This method needs further investigation before any statement about clinical value can be made.

Another method that is still in its infancy is aptamer-based CTC detection. Aptamers are small nucleotide sequences that can be designed against any molecular target and can bind to it with high specificity and affinity. Zamay at al. (130) were able to select and identify DNA aptamers for lung cancer cells from surgical excision tissue. These aptamers were then used to identify CTCs from patient blood plated onto slides after RBC lysis. Of 52 primary lung cancer patients, 44 were positive for CTCs, and of that subset, 15 of those had less than 2 cells/3 mL.

ScreenCell® (ScreenCell SA, Paris, France) is a newer filtration device with choices of pore sizes (8 or 6.5 μm) and buffers for either fixation or cell culture. It is possible to filter CTCs and directly culture them for further experiments. Recovery of cell line samples has been demonstrated at approximately 90%, with more than 85% of the sample being viable cells (131), suggesting that ScreenCell® may be a useful combination method by which to obtain living CTC cultures or xenograft models.

A big step toward automation was achieved by the fully automated DEPArray™ (Silicon Biosystems, San Diego, CA) instrument, which detects, quantifies, and recovers CTCs for downstream analysis from CTC-enriched blood samples. This method needs a CTC enrichment step like CELLSEARCH® or OncoQuick® before further analysis, but subsequently everything is automated, with minimal hands-on requirements. The enriched cell solution is pipetted into a chip and is loaded to the DEPArray™ instrument. Within the chip are various microelectrodes, which produce electric cages to trap (up to 40,000) single cells. Once the cells are trapped in individual cages, they can be identified and selected on the basis of fluorescent markers. Single cells can be moved within the chip by electrical forces and recovered in a tube for further genomic analysis (132, 133). Fernandez et al. (132) reported CTC subclones in metastatic breast cancer patients; some patients showed genomic concordance with the primary tumor, and some did not. These highly automated techniques could therefore be a useful tool to monitor disease progression and tumor heterogeneity.

Nonenrichment Strategies for Detection: Direct Analysis

The HD-SCA assay developed by Kuhn and colleagues (10) is under commercialization by Epic Sciences (San Diego, CA). This technique uses a more holistic approach to CTC detection. Peripheral blood is drawn, and within 48 h, RBC lysis is performed, and peripheral blood mononuclear cells (PBMCs) are plated on customized glass slides and stained for HD-CTC identification with antibodies against a CK-pan mix, CD45, and a nuclei counterstain with DAPI. For image analysis, an automated system is used, and custom-made software generates a report that presents the candidate cells to a hematopathologist-trained technical analyst for analysis and interpretation. This approach not only allows nucleated CK-positive, CD45-negative cells to be reported in CTC analysis, but also records all cells of interest as shown in **Figure 1**. No subpopulation is selected for, as in other methods described, which use surface markers or other limited properties for detection. All cells of the PBMC fraction of each patient are recorded and imaged, and no image is discarded. In a study of 28 small cell lung cancer patients, Kuhn and colleagues (134) detected numbers as low as 0.3 CTCs/mL at study enrollment. This number increased to an average of 13.4 CTCs/mL, with higher CTC counts in follow-up samples. This platform recently underwent major enhancement by integrating NGS at the single-cell level. Any CTC detected in a patient sample can be picked from its slide and can undergo CNV analysis or mutation screening to uncover the clonal relationships within and between CTC populations and the primary tumor. A drawback of this direct-analysis platform is the fixation of all cells such that no viable CTCs can be obtained for further studies.

DOWNSTREAM CHARACTERIZATION OF CIRCULATING TUMOR CELLS

Establishment of Circulating Tumor Cell Cultures and Xenograft Models

One highly attractive downstream option for further investigation of CTCs after they are identified is the culturing of viable CTCs. Due to the scarcity of CTCs, this approach cannot be implemented

for all patient samples. For example, in cases of prostate cancer, at least 100 cells are required for successful culture of CTCs (135). Once a CTC cell line is established, it can either be used for direct testing of drug sensitivity (47, 136) or be implanted into immunosuppressed mice to create xenograft models and thus permit further drug testing and genetic profiling studies (137). However, either option limits the ability to study fundamental heterogeneity, as only the most robust cells are likely to survive the process of ex vivo engraftment. Moreover, the cells will have been removed from their host environment and immune system, which may alter their nature, precluding meaningful conclusions regarding disease evolution or treatment resistance (138).

Molecular Characterization of Circulating Tumor Cells

The recent prevailing aim is characterization of the DNA content of single CTCs. Over the past decade, NGS methods have immensely increased our knowledge about cancer genotypes and their derived CTCs (139, 140). Performing any single-cell method at the DNA level first requires whole-genome amplification (WGA) to increase the minute amount of 6.6 pg total DNA/cell before further analysis (42). This process increases copy numbers of the entire genome but increases the risk of bias and thus false findings (141). Therefore, studies have to include a comparison of CTCs to normal cells or in some other way demonstrate fidelity during amplification. After WGA the genome can be analyzed in multiple ways: mutation analysis of therapeutic target genes (42, 142), massive parallel sequencing for the detection of new druggable mutations (143), or CNV detection through NGS (144). Mutations in *KRAS*, for example, can block EGFR-targeted therapy efficiency in colorectal cancer. Studies revealed high intrapatient and interpatient heterogeneity of *KRAS* mutations (143). One theory is that CTCs carrying the mutation may escape therapy and then cause relapse or disease progression. The ability to detect and identify these cells early could lead to adjustments in precision medicine (22).

RNA-based CTC profiling also offers promise. Gene expression studies at the RNA level may reveal important information about tumor heterogeneity. Popular techniques are fluorescence in situ hybridization (FISH) (145), reverse transcription PCR followed by other PCR techniques (like quantitative real-time PCR) (42), and microarray mRNA sequencing (115). Using quantitative real-time PCR, Steinestel et al. (146) investigated the two major mutations—the *AR-V7* and *AR* point mutations—in the androgen receptor gene (*AR*) in advanced prostate cancer CTCs in the context of probable drug resistance. Patients underwent a molecularly uninformed therapy switch with a response rate of 38%. Steinestel and colleagues then discovered in retrospect that 71.4% of the responders had had CTCs that showed a matching molecular profile for the therapy switch. The overall calculated benefits of known molecular profiles of CTC for the success of a therapy switch were estimated to be ~27%.

Finally, CTCs can be interrogated at the protein level. The most common approach by which to attain expression data from CTCs is protein detection through direct antibody-target binding. This method has become standard and is integrated within many techniques like CELLSEARCH® (41). These techniques usually evaluate the hormone status of patients and compare protein expression levels of the tissue sample with CTC expression rates. Examples of proteins of interest are HER2 in breast cancer (41, 147) and PSMA in prostate cancer (148). Studies are still ongoing to determine whether patients can benefit from a therapy switch based on CTC marker expression. Techniques using antibody-target binding include immunohistochemistry approaches using either labeled fluorophores or magnetic particles with subsequent microscopy (10, 149) and enzyme-linked immunosorbent assay (150). Improved high-throughput methods include protein microarrays (which have a device surface covered with capture antibodies) and mass spectrometry,

both of which offer high sensitivity and specificity (151). Additionally, nanotechnology can be used to detect protein expression in CTCs through labeled nanoparticles or nanowires (116).

CLINICAL IMPLICATIONS

Using these new technologies to fully interrogate CTCs could lead to many possible clinical applications. Studies demonstrating the predictive value of CTCs for OS or PFS in cancer patients have been copious over the last decade (79, 110, 150, 152). However, CTCs carry much greater promise than has yet been realized, and within the last three years the focus in CTC research has shifted to studies focusing on the transfer of research results to clinical practice (138).

Therapeutic Target Discovery

CTCs may allow for the detection of known therapeutic targets in situations in which tissue is limited. An excellent example is lung cancer. In NSCLC the diagnostic test for crizotinib treatment effectiveness in rearranged anaplastic lymphoma kinase (ALK) patients is currently performed on tumor biopsy samples. To determine whether CTCs could provide the same diagnostic information—which could be especially useful when biopsy tissue is limited, as it so often is for fine-needle lung biopsies—Pailler et al. (153) designed a platform for CTCs, using a filter-adapted FISH method in combination with the FDA-approved companion diagnostic FISH probe. They used a NSCLC cohort of 32 patients and compared results from CTCs with those from tumor biopsies. This study produced concordant results of CTCs with biopsy examinations for *ALK* rearrangements. All ALK-positive patients had CTCs with *ALK* rearrangements, and there was only one *ALK* rearrangement in CTCs of the ALK-negative patients. Interestingly, *ALK*-rearranged CTCs were positive for EMT markers like vimentin and N-cadherin but had no CK; the authors suggest that these cells thus showed a more mesenchymal phenotype, which implies a higher metastatic potential. The group later showed the system to be effective with c-*ros oncogene 1* as well (154).

Disease and Treatment Monitoring

CTCs may provide markers for treatment sensitivity. In 62 prostate cancer patients, men with AR-V7-positive CTCs had shorter PFS and OS than did AR-V7-negative patients, which shows that AR-V7-positive CTCs predict resistance to enzalutamide and abiraterone (155). One year after this study, the group investigated the same marker in correlation with taxane treatment and observed that taxane is more effective in men with AR-V7-positive CTCs and shows an efficiency comparable to that of enzalutamide and abiraterone in AR-V7-negative patients (156).

CTCs may predict treatment response. For example, Wallwiener et al. (157) demonstrated that in metastatic breast cancer a continuously high CTC level after one cycle of chemotherapy predicts shorter OS and PFS and that CTCs therefore may be useful in adjusting systemic therapies. The SWOG S0500 trial confirmed CTC counts to be a predictive marker for PFS and OS but failed to show improved survival after a therapy switch based on CTC counts (158). The Wallwiener group stated that these results could also indicate that there is a need for more effective treatment for this subpopulation of patients (157).

CTCs may allow us to recognize heterogeneity and better investigate therapy resistance mechanisms and to even identify novel therapeutic targets. NGS has revealed that tumor heterogeneity is dynamic, with biomarkers changing during disease progression (4). CTCs may provide an even more precise representation of the mutation spectrum of the primary tumor. This theory is supported by findings of Heitzer et al. (143), who investigated concordance between the primary

tumor, the metastasis, and CTCs for KRAS. Driver mutations like those in the genes encoding KRAS, APC, and PIK3CA were found mostly in metastases and the corresponding CTCs, but some mutations appeared only in CTCs. This theory could potentially explain some cases of treatment resistance seen in, for example, hormone receptor–positive patients who show no benefit from endocrine treatment. Changes in estrogen receptor and/or progesterone receptor expression between solid tumor sites and CTCs may explain the therapy failure (3). CTCs that derive from an overall marker-positive site may in fact be heterogeneous, and the important subpopulations that survive treatment and cause disease progression or recurrence may for this reason be found among the CTCs.

Prediction of Risk of Relapse

CTCs may be prognostic of relapse. A large multicenter study was conducted with 2,026 early breast cancer patients by using the CELLSEARCH® system. The worst prognosis was shown for patients with ≥5 CTCs/30 mL of blood, and the detection of CTCs before and after chemotherapy was linked to an increased risk of relapse (12). In addition to this breast cancer study, studies of prostate cancer (159) and bladder cancer (160) have shown a predictive correlation of CTC counts with metastatic relapse.

Technologies incorporating downstream analysis and characterization of CTCs after detection can be applied to determine the metastatic potential of CTCs by characterizing mutations, CNVs, or protein expression changes in one or more of the factors involved in the metastatic cascade. Baccelli and colleagues (124) characterized the metastatic potential of CTC subpopulations. The presence of a group of EpCAM-positive, CD44-positive, CD47-positive, and MET-positive CTCs was correlated with patient survival and metastasis. This identification of metastasis-initiating CTCs and the markers used could play a key role in prognosis of relapse or metastasis. Given that CTCs can be detected before metastasis, their analysis may allow for more precise staging of early-stage patients and thus function as a prognostic tool for therapy decision making—one that can easily, reproducibly, and repeatedly be obtained through a fluid biopsy from cancer patients (161).

Treatment Vehicles

Finally, at the very cutting edge, CTCs have been proposed as treatment vehicles in a potential Trojan horse role. Targeting CTCs with genetically modified platelets that express TRAIL (tumor necrosis factor–related apoptosis-inducing ligand) could induce apoptosis in tumor cells (162). Due to platelet interactions in the bloodstream, CTCs would be the first target of these modified platelets, as successfully shown in vitro and in mouse models. This strategy could be a next step toward slowing down the metastatic cascade (162).

DISCLOSURE STATEMENT

P.K. is a shareholder of and consultant to Epic Sciences.

ACKNOWLEDGMENTS

This work was supported by National Sustainability Program I LO1503 provided by the Ministry of Education, Youth, and Sports of the Czech Republic. We thank Michaela Brychtová, MUDr., PhD, for preparing **Figure 3**.

LITERATURE CITED

1. Siegel RL, Miller KD, Jemal A. 2015. Cancer statistics, 2015. *CA Cancer J. Clin.* 65(1):5–29
2. Burrell RA, McGranahan N, Bartek J, Swanton C. 2013. The causes and consequences of genetic heterogeneity in cancer evolution. *Nature* 501(7467):338–45
3. Babayan A, Hannemann J, Spötter J, Müller V, Pantel K, Joosse SA. 2013. Heterogeneity of estrogen receptor expression in circulating tumor cells from metastatic breast cancer patients. *PLOS ONE* 8(9):e75038
4. Zhang C, Guan Y, Sun Y, Ai D, Guo Q. 2016. Tumor heterogeneity and circulating tumor cells. *Cancer Lett.* 374(2):216–23
5. de Bruin EC, McGranahan N, Mitter R, Salm M, Wedge DC, et al. 2014. Spatial and temporal diversity in genomic instability processes defines lung cancer evolution. *Science* 346(6206):251–56
6. Ashworth TR. 1869. A case of cancer in which cells similar to those in the tumours were seen in the blood after death. *Med. J. Aust.* 14:146–47
7. Fischer AH. 2009. Circulating tumor cells: Seeing is believing. *Arch. Pathol. Lab. Med.* 133(9):1367–69
8. Miller MC, Doyle GV, Terstappen LWMM. 2010. Significance of circulating tumor cells detected by the CellSearch system in patients with metastatic breast colorectal and prostate cancer. *J. Oncol.* 2010:617421
9. Allard WJ, Matera J, Miller MC, Repollet M, Connelly MC, et al. 2004. Tumor cells circulate in the peripheral blood of all major carcinomas but not in healthy subjects or patients with nonmalignant diseases. *Clin. Cancer Res.* 10(20):6897–904
10. Marrinucci D, Bethel K, Kolatkar A, Luttgen MS, Malchiodi M, et al. 2012. Fluid biopsy in patients with metastatic prostate, pancreatic and breast cancers. *Phys. Biol.* 9(1):016003
11. Nagrath S, Sequist LV, Maheswaran S, Bell DW, Irimia D, et al. 2007. Isolation of rare circulating tumour cells in cancer patients by microchip technology. *Nature* 450(7173):1235–39
12. Rack B, Schindlbeck C, Jückstock J, Andergassen U, Hepp P, et al. 2014. Circulating tumor cells predict survival in early average-to-high risk breast cancer patients. *J. Natl. Cancer Inst.* 106(5):dju066
13. Balic M, Dandachi N, Hofmann G, Samonigg H, Loibner H, et al. 2005. Comparison of two methods for enumerating circulating tumor cells in carcinoma patients. *Cytom. B Clin. Cytom.* 68(1):25–30
14. Riethdorf S, Fritsche H, Muller V, Rau T, Schindlbeck C, et al. 2007. Detection of circulating tumor cells in peripheral blood of patients with metastatic breast cancer: a validation study of the CellSearch system. *Clin. Cancer Res.* 13(3):920–28
15. Drye JC, Rumage WT, Anderson D. 1962. Prognostic import of circulating cancer cells after curative surgery: a long time follow up study. *Ann. Surg.* 155:733–40
16. Salsbury AJ. 1975. The significance of the circulating cancer cell. *Cancer Treat. Rev.* 2(1):55–72
17. World Health Organ. (WHO). 1963. *Cancer control: first report of an expert committee*. Tech. Rep. 251, WHO
18. Bao H, Burke PA, Huang J, Chen X, Brohawn PZ, et al. 2013. Circulating tumor cells: application as a biomarker for molecular characterization and predictor of survival in an all-comer solid tumor Phase I clinical study. *PLOS ONE* 8(8):e58557
19. Gradishar WJ, Anderson BO, Balassanian R, Blair SL, Burstein HJ, et al. 2015. Breast Cancer Version 2.2015: clinical practice guidelines in oncology. *J. Natl. Compr. Cancer Netw.* 13(4):448–75
20. Ruiz C, Li J, Luttgen MS, Kolatkar A, Kendall JT, et al. 2015. Limited genomic heterogeneity of circulating melanoma cells in advanced stage patients. *Phys. Biol.* 12(1):016008
21. Chang L, Asatrian G, Dry SM, James AW. 2015. Circulating tumor cells in sarcomas: a brief review. *Med. Oncol.* 32(1):430
22. Alix-Panabières C, Pantel K. 2016. Clinical applications of circulating tumor cells and circulating tumor DNA as liquid biopsy. *Cancer Discov.* 6:479–91
23. Pantel K, Alix-Panabières C. 2016. Liquid biopsy: potential and challenges. *Mol. Oncol.* 10(3):371–73
24. Labelle M, Begum S, Hynes RO. 2011. Direct signaling between platelets and cancer cells induces an epithelial-mesenchymal-like transition and promotes metastasis. *Cancer Cell* 20(5):576–90
25. Gold B, Cankovic M, Furtado LV, Meier F, Gocke CD. 2015. Do circulating tumor cells, exosomes, and circulating tumor nucleic acids have clinical utility? A report of the Association for Molecular Pathology. *J. Mol. Diagn.* 17(3):209–24

26. Newman AM, Bratman SV, To J, Wynne JF, Eclov NCW, et al. 2014. An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage. *Nat. Med.* 20(5):548–54
27. Harouaka RA, Nisic M, Zheng S-Y. 2013. Circulating tumor cell enrichment based on physical properties. *J. Lab. Autom.* 18(6):455–68
28. Shapiro HM, Schildkraut ER, Curbelo R, Laird CW, Turner B, Hirschfeld T. 1976. Combined blood cell counting and classification with fluorochrome stains and flow instrumentation. *J. Histochem. Cytochem.* 24(1):396–401
29. Guck J, Schinkinger S, Lincoln B, Wottawah F, Ebert S, et al. 2005. Optical deformability as an inherent cell marker for testing malignant transformation and metastatic competence. *Biophys. J.* 88(5):3689–98
30. Suresh S. 2007. Biomechanics and biophysics of cancer cells. *Acta Biomater.* 3(4):413–38
31. Park ES, Jin C, Guo Q, Ang RR, Duffy SP, et al. 2016. Continuous flow deformability-based separation of circulating tumor cells using microfluidic ratchets. *Small* 12(14):1909–19
32. Coughlin MF, Bielenberg DR, Lenormand G, Marinkovic M, Waghorne CG, et al. 2013. Cytoskeletal stiffness, friction, and fluidity of cancer cell lines with different metastatic potential. *Clin. Exp. Metastasis* 30(3):237–50
33. Cross SE, Jin Y-S, Rao J, Gimzewski JK. 2007. Nanomechanical analysis of cells from cancer patients. *Nat. Nanotechnol.* 2(12):780–83
34. Shaw Bagnall J, Byun S, Begum S, Miyamoto DT, Hecht VC, et al. 2015. Deformability of tumor cells versus blood cells. *Sci. Rep.* 5:18542
35. Pethig R. 2010. Dielectrophoresis: status of the theory, technology, and applications. *Biomicrofluidics* 4(2):022811
36. Gascoyne P, Shim S. 2014. Isolation of circulating tumor cells by dielectrophoresis. *Cancers* 6(1):545–79
37. Davis D, Gupta V, Garza M, Pace M, Wu W, et al. 2011. EpCAM-independent ApoStream^{TM} technology isolates circulating tumor cells from blood of patients with various types of cancer. *Mol. Cancer Ther.* 10(Suppl. 1):B20 (Abstr.)
38. Ponder E. 1942. The relation between red blood cell density and corpuscular hemoglobin concentration. *J. Biol. Chem.* 144:333–38
39. Phillips KG, Velasco CR, Li J, Kolatkar A, Luttgen M, et al. 2012. Optical quantification of cellular mass, volume, and density of circulating tumor cells identified in an ovarian cancer patient. *Front. Oncol.* 2:72
40. Nadal R, Lorente JA, Rosell R, Serrano MJ. 2013. Relevance of molecular characterization of circulating tumor cells in breast cancer in the era of targeted therapies. *Expert Rev. Mol. Diagn.* 13(3):295–307
41. Fehm T, Müller V, Aktas B, Janni W, Schneeweiss A, et al. 2010. Her2 status of circulating tumor cells in patients with metastatic breast cancer: a prospective, multicenter trial. *Breast Cancer Res. Treat.* 124(2):403–12
42. Gasch C, Bauernhofer T, Pichler M, Langer-Freitag S, Reeh M, et al. 2013. Heterogeneity of epidermal growth factor receptor status and mutations of *KRAS/PIK3CA* in circulating tumor cells of patients with colorectal cancer. *Clin. Chem.* 59(1):252–60
43. Schramm A, Friedl TWP, Schochter F, Scholz C, de Gregorio N, et al. 2016. Therapeutic intervention based on circulating tumor cell phenotype in metastatic breast cancer: concept of the DETECT study program. *Arch. Gynecol. Obstet.* 293(2):271–81
44. Gazzaniga P, Raimondi C, Gradilone A, Cortesi E, Naso G. 2014. Clinical utility of circulating tumor cell counting through CellSearch®: the dilemma of a concept suspended in limbo. *OncoTargets Ther.* 7:619–25
45. Masuda T, Hayashi N, Iguchi T, Ito S, Eguchi H, Mimori K. 2016. Clinical and biological significance of circulating tumor cells in cancer. *Mol. Oncol.* 10(3):408–17
46. Marrinucci D, Bethel K, Bruce RH, Curry DN, Hsieh B, et al. 2007. Case study of the morphologic variation of circulating tumor cells. *Hum. Pathol.* 38(3):514–19
47. Kolostova K, Spicka J, Matkowski R, Bobek V. 2015. Isolation, primary culture, morphological and molecular characterization of circulating tumor cells in gynecological cancers. *Am. J. Transl. Res.* 7(7):1203–13

48. Jahr S, Hentze H, Englisch S, Hardt D, Fackelmayer FO, et al. 2001. DNA fragments in the blood plasma of cancer patients: quantitations and evidence for their origin from apoptotic and necrotic cells. *Cancer Res.* 61(4):1659–65
49. Krebs MG, Metcalf RL, Carter L, Brady G, Blackhall FH, Dive C. 2014. Molecular analysis of circulating tumour cells—biology and biomarkers. *Nat. Rev. Clin. Oncol.* 11(3):129–44
50. Park S, Ang RR, Duffy SP, Bazov J, Chi KN, et al. 2014. Morphological differences between circulating tumor cells from prostate cancer patients and cultured prostate cancer cells. *PLOS ONE* 9(1):e85264
51. Ligthart ST, Coumans FAW, Bidard F-C, Simkens LHJ, Punt CJA, et al. 2013. Circulating tumor cells count and morphological features in breast, colorectal and prostate cancer. *PLOS ONE* 8(6):e67148
52. Pantel K, Speicher MR. 2016. The biology of circulating tumor cells. *Oncogene* 35(10):1216–24
53. Cavallaro U, Christofori G. 2001. Cell adhesion in tumor invasion and metastasis: Loss of the glue is not enough. *Biochim. Biophys. Acta* 1552(1):39–45
54. Liotta LA, Kleinerman J, Saidel GM. 1974. Quantitative relationships of intravascular tumor cells, tumor vessels, and pulmonary metastases following tumor implantation. *Cancer Res.* 34(5):997–1004
55. Chang YS, di Tomaso E, McDonald DM, Jones R, Jain RK, Munn LL. 2000. Mosaic blood vessels in tumors: frequency of cancer cells in contact with flowing blood. *PNAS* 97(26):14608–13
56. Condeelis J, Pollard JW. 2006. Macrophages: obligate partners for tumor cell migration, invasion, and metastasis. *Cell* 124(2):263–66
57. Bockhorn M, Jain RK, Munn LL. 2007. Active versus passive mechanisms in metastasis: Do cancer cells crawl into vessels, or are they pushed? *Lancet Oncol.* 8(5):444–48
58. Kalluri R, Weinberg RA. 2009. The basics of epithelial-mesenchymal transition. *J. Clin. Investig.* 119(6):1420–28
59. Pietilä M, Ivaska J, Mani SA. 2016. Whom to blame for metastasis, the epithelial-mesenchymal transition or the tumor microenvironment? *Cancer Lett.* 380(1):359–68
60. Laberge R-M, Awad P, Campisi J, Desprez P-Y. 2012. Epithelial-mesenchymal transition induced by senescent fibroblasts. *Cancer Microenviron.* 5(1):39–44
61. Liu Z-J, Semenza GL, Zhang H-F. 2015. Hypoxia-inducible factor 1 and breast cancer metastasis. *J. Zhejiang Univ. Sci. B* 16(1):32–43
62. Bacigalupo ML, Manzi M, Espelt MV, Gentilini LD, Compagno D, et al. 2015. Galectin-1 triggers epithelial-mesenchymal transition in human hepatocellular carcinoma cells. *J. Cell. Physiol.* 230(6):1298–309
63. Mamuya FA, Duncan MK. 2012. αV integrins and TGF-β-induced EMT: a circle of regulation. *J. Cell. Mol. Med.* 16(3):445–55
64. Hazan RB, Qiao R, Keren R, Badano I, Suyama K. 2004. Cadherin switch in tumor progression. *Ann. N. Y. Acad. Sci.* 1014:155–63
65. Martin TA, Ye L, Sanders AJ, Lane J, Jiang WG. 2013. Cancer invasion and metastasis: molecular and cellular perspective. In *Metastatic Cancer: Clinical and Biological Perspectives*, ed. R Jandial. Austin: Landes Biosci.
66. Chang JT, Mani SA. 2013. Sheep, wolf, or werewolf: cancer stem cells and the epithelial-to-mesenchymal transition. *Cancer Lett.* 341(1):16–23
67. Shibue T, Brooks MW, Weinberg RA. 2013. An integrin-linked machinery of cytoskeletal regulation that enables experimental tumor initiation and metastatic colonization. *Cancer Cell* 24(4):481–98
68. Sleeman JP, Thiery JP. 2011. SnapShot: the epithelial-mesenchymal transition. *Cell* 145(1):162.e1
69. Clark AG, Vignjevic DM. 2015. Modes of cancer cell invasion and the role of the microenvironment. *Curr. Opin. Cell Biol.* 36:13–22
70. Ueno H, Shinto E, Kajiwara Y, Fukazawa S, Shimazaki H, et al. 2014. Prognostic impact of histological categorisation of epithelial-mesenchymal transition in colorectal cancer. *Br. J. Cancer* 111(11):2082–90
71. Kohler I, Bronsert P, Timme S, Werner M, Brabletz T, et al. 2015. Detailed analysis of epithelial-mesenchymal transition and tumor budding identifies predictors of long-term survival in pancreatic ductal adenocarcinoma. *J. Gastroenterol. Hepatol.* 30(Suppl. 1):78–84
72. Stewart CJR, McCluggage WG. 2013. Epithelial-mesenchymal transition in carcinomas of the female genital tract. *Histopathology* 62(1):31–43

73. Tsuji T, Ibaragi S, Hu G-F. 2009. Epithelial-mesenchymal transition and cell cooperativity in metastasis. *Cancer Res.* 69(18):7135–39
74. Aceto N, Bardia A, Miyamoto DT, Donaldson MC, Wittner BS, et al. 2014. Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis. *Cell* 158(5):1110–22
75. Fischer KR, Durrans A, Lee S, Sheng J, Li F, et al. 2015. Epithelial-to-mesenchymal transition is not required for lung metastasis but contributes to chemoresistance. *Nature* 527(7579):472–76
76. Zheng X, Carstens JL, Kim J, Scheible M, Kaye J, et al. 2015. Epithelial-to-mesenchymal transition is dispensable for metastasis but induces chemoresistance in pancreatic cancer. *Nature* 527(7579):525–30
77. Frisch SM, Francis H. 1994. Disruption of epithelial cell-matrix interactions induces apoptosis. *J. Cell Biol.* 124(4):619–26
78. Bockhorn M, Roberge S, Sousa C, Jain RK, Munn LL. 2004. Differential gene expression in metastasizing cells shed from kidney tumors. *Cancer Res.* 64(7):2469–73
79. Meng S, Tripathy D, Frenkel EP, Shete S, Naftalis EZ, et al. 2004. Circulating tumor cells in patients with breast cancer dormancy. *Clin. Cancer Res.* 10:8152–62
80. Rossi E, Basso U, Celadin R, Zilio F, Pucciarelli S, et al. 2010. M30 neoepitope expression in epithelial cancer: quantification of apoptosis in circulating tumor cells by CellSearch analysis. *Clin. Cancer Res.* 16(21):5233–43
81. Cho EH, Wendel M, Luttgen M, Yoshioka C, Marrinucci D, et al. 2012. Characterization of circulating tumor cell aggregates identified in patients with epithelial tumors. *Phys. Biol.* 9(1):016001
82. Kats-Ugurlu G, Roodink I, de Weijert M, Tiemessen D, Maass C, et al. 2009. Circulating tumour tissue fragments in patients with pulmonary metastasis of clear cell renal cell carcinoma. *J. Pathol.* 219(3):287–93
83. Khorana AA, Francis CW, Culakova E, Kuderer NM, Lyman GH. 2007. Thromboembolism is a leading cause of death in cancer patients receiving outpatient chemotherapy. *J. Thromb. Haemost.* 5(3):632–34
84. Shiao SL, Chu GC-Y, Chung LWK. 2016. Regulation of prostate cancer progression by the tumor microenvironment. *Cancer Lett.* 380(1):340–48
85. Zhang J, Qiao X, Shi H, Han X, Liu W, et al. 2016. Circulating tumor-associated neutrophils (cTAN) contribute to circulating tumor cell survival by suppressing peripheral leukocyte activation. *Tumour Biol.* 37(4):5397–404
86. Yadav A, Kumar B, Yu J-G, Old M, Teknos TN, Kumar P. 2015. Tumor-associated endothelial cells promote tumor metastasis by chaperoning circulating tumor cells and protecting them from anoikis. *PLOS ONE* 10(10):e0141602
87. Liotta LA, Saidel MG, Kleinerman J. 1976. The significance of hematogenous tumor cell clumps in the metastatic process. *Cancer Res.* 36(3):889–94
88. Hou J-M, Krebs MG, Lancashire L, Sloane R, Backen A, et al. 2012. Clinical significance and molecular characteristics of circulating tumor cells and circulating tumor microemboli in patients with small-cell lung cancer. *J. Clin. Oncol.* 30(5):525–32
89. Yu JL, May L, Lhotak V, Shahrzad S, Shirasawa S, et al. 2005. Oncogenic events regulate tissue factor expression in colorectal cancer cells: implications for tumor progression and angiogenesis. *Blood* 105(4):1734–41
90. Phillips KG, Lee AM, Tormoen GW, Rigg RA, Kolatkar A, et al. 2015. The thrombotic potential of circulating tumor microemboli: computational modeling of circulating tumor cell–induced coagulation. *Am. J. Physiol. Cell Physiol.* 308(3):C229–36
91. Mitrugno A, Tormoen GW, Kuhn P, McCarty OJT. 2016. The prothrombotic activity of cancer cells in the circulation. *Blood Rev.* 30(1):11–19
92. Zhe X, Cher ML, Bonfil RD. 2011. Circulating tumor cells: finding the needle in the haystack. *Am. J. Cancer Res.* 1(6):740–51
93. Xu L, Shamash J, Lu Y-J. 2015. Circulating tumor cells: a window to understand cancer metastasis, monitor and fight against cancers. *J. Cancer Res. Updat.* 4(1):13–29
94. Duda DG, Duyverman AMMJ, Kohno M, Snuderl M, Steller EJA, et al. 2010. Malignant cells facilitate lung metastasis by bringing their own soil. *PNAS* 107(50):21677–82
95. Norton L, Massagué J. 2006. Is cancer a disease of self-seeding? *Nat. Med.* 12(8):875–78
96. Paget S. 1989. The distribution of secondary growths in cancer of the breast. *Cancer Metastasis Rev.* 8(2):98–101

97. Comen E, Norton L, Massagué J. 2011. Clinical implications of cancer self-seeding. *Nat. Rev. Clin. Oncol.* 8(6):369–77
98. Pattabiraman DR, Bierie B, Kober KI, Thiru P, Krall JA, et al. 2016. Activation of PKA leads to mesenchymal-to-epithelial transition and loss of tumor-initiating ability. *Science* 351(6277):aad3680
99. Yao D, Dai C, Peng S. 2011. Mechanism of the mesenchymal-epithelial transition and its relationship with metastatic tumor formation. *Mol. Cancer Res.* 9(12):1608–20
100. Massagué J, Obenauf AC. 2016. Metastatic colonization by circulating tumour cells. *Nature* 529(7586):298–306
101. Kim M-Y, Oskarsson T, Acharyya S, Nguyen DX, Zhang XH-F, et al. 2009. Tumor self-seeding by circulating cancer cells. *Cell* 139(7):1315–26
102. Newton PK, Mason J, Bethel K, Bazhenova L, Nieva J, et al. 2013. Spreaders and sponges define metastasis in lung cancer: a Markov chain Monte Carlo mathematical model. *Cancer Res.* 73(9):2760–69
103. Chaffer CL, Weinberg RA. 2011. A perspective on cancer cell metastasis. *Science* 331(6024):1559–64
104. Spiliotaki M, Mavroudis D, Kapranou K, Markomanolaki H, Kallergi G, et al. 2014. Evaluation of proliferation and apoptosis markers in circulating tumor cells of women with early breast cancer who are candidates for tumor dormancy. *Breast Cancer Res.* 16(6):485
105. McGovern M, Voutev R, Maciejowski J, Corsi AK, Hubbard EJA. 2009. A "latent niche" mechanism for tumor initiation. *PNAS* 106(28):11617–22
106. Yeh AC, Ramaswamy S. 2015. Mechanisms of cancer cell dormancy—another hallmark of cancer? *Cancer Res.* 75(23):5014–22
107. Joosse SA, Gorges TM, Pantel K. 2015. Biology, detection, and clinical implications of circulating tumor cells. *EMBO Mol. Med.* 7(1):1–11
108. Alix-Panabières C, Pantel K. 2014. Challenges in circulating tumour cell research. *Nat. Rev. Cancer* 14(9):623–31
109. Harouaka R, Kang Z, Zheng S-Y, Cao L. 2014. Circulating tumor cells: advances in isolation and analysis, and challenges for clinical applications. *Pharmacol. Ther.* 141(2):209–21
110. Cristofanilli M, Hayes DF, Budd GT, Ellis MJ, Stopeck A, et al. 2005. Circulating tumor cells: a novel prognostic factor for newly diagnosed metastatic breast cancer. *J. Clin. Oncol.* 23(7):1420–30
111. Stott SL, Hsu C-H, Tsukrov DI, Yu M, Miyamoto DT, et al. 2010. Isolation of circulating tumor cells using a microvortex-generating herringbone-chip. *PNAS* 107(43):18392–97
112. Wang S, Thomas A, Lee E, Yang S, Cheng X, Liu Y. 2016. Highly efficient and selective isolation of rare tumor cells using a microfluidic chip with wavy-herringbone micro-patterned surfaces. *Analyst* 141:2228–37
113. Wang S, Liu K, Liu J, Yu ZT-F, Xu X, et al. 2011. Highly efficient capture of circulating tumor cells by using nanostructured silicon substrates with integrated chaotic micromixers. *Angew. Chem. Int. Ed.* 50(13):3084–88
114. Talasaz AH, Powell AA, Huber DE, Berbee JG, Roh K-H, et al. 2009. Isolating highly enriched populations of circulating epithelial cells and other rare cells from blood using a magnetic sweeper device. *PNAS* 106(10):3970–75
115. Cann GM, Gulzar ZG, Cooper S, Li R, Luo S, et al. 2012. mRNA-Seq of single prostate cancer circulating tumor cells reveals recapitulation of gene expression and pathways found in prostate cancer. *PLOS ONE* 7(11):e49144
116. Saucedo-Zeni N, Mewes S, Niestroj R, Gasiorowski L, Murawa D, et al. 2012. A novel method for the in vivo isolation of circulating tumor cells from peripheral blood of cancer patients using a functionalized and structured medical wire. *Int. J. Oncol.* 41(4):1241–50
117. Gorges TM, Penkalla N, Schalk T, Joosse SA, Riethdorf S, et al. 2016. Enumeration and molecular characterization of tumor cells in lung cancer patients using a novel in vivo device for capturing circulating tumor cells. *Clin. Cancer Res.* 22(9):2197–206
118. Gorges TM, Tinhofer I, Drosch M, Röse L, Zollner TM, et al. 2012. Circulating tumour cells escape from EpCAM-based detection due to epithelial-to-mesenchymal transition. *BMC Cancer* 12:178
119. Baker MK, Mikhitarian K, Osta W, Callahan K, Hoda R, et al. 2003. Molecular detection of breast cancer cells in the peripheral blood of advanced-stage breast cancer patients using multimarker real-time reverse transcription-polymerase chain reaction and a novel porous barrier density gradient centrifugation technology. *Clin. Cancer Res.* 9(13):4865–71

120. Chinen LTD, de Carvalho FM, Rocha BMM, Aguiar CM, Abdallah EA, et al. 2013. Cytokeratin-based CTC counting unrelated to clinical follow up. *J. Thorac. Dis.* 5(5):593–99
121. Tan SJ, Yobas L, Lee GYH, Ong CN, Lim CT. 2009. Microdevice for the isolation and enumeration of cancer cells from blood. *Biomed. Microdevices* 11(4):883–92
122. Hou HW, Warkiani ME, Khoo BL, Li ZR, Soo RA, et al. 2013. Isolation and retrieval of circulating tumor cells using centrifugal forces. *Sci. Rep.* 3:1259
123. Yin J, Wang Y, Yin H, Chen W, Jin G, et al. 2015. Circulating tumor cells enriched by the depletion of leukocytes with bi-antibodies in non–small cell lung cancer: potential clinical application. *PLOS ONE* 10(8):e0137076
124. Baccelli I, Schneeweiss A, Riethdorf S, Stenzinger A, Schillert A, et al. 2013. Identification of a population of blood circulating tumor cells from breast cancer patients that initiates metastasis in a xenograft assay. *Nat. Biotechnol.* 31(6):539–44
125. Alix-Panabières C. 2012. EPISPOT assay: detection of viable DTCs/CTCs in solid tumor patients. In *Minimal Residual Disease and Circulating Tumor Cells in Breast Cancer. Recent Results in Cancer Research, Volume 195*, ed. M Ignatiadis, C Sotiriou, K Pantel, pp. 69–76. Berlin/Heidelberg, Ger.: Springer
126. Ramirez J-M, Fehm T, Orsini M, Cayrefourcq L, Maudelonde T, et al. 2014. Prognostic relevance of viable circulating tumor cells detected by EPISPOT in metastatic breast cancer patients. *Clin. Chem.* 60(1):214–21
127. Pizon M, Zimon D, Carl S, Pachmann U, Pachmann K, Camara O. 2013. Heterogeneity of circulating epithelial tumour cells from individual patients with respect to expression profiles and clonal growth (sphere formation) in breast cancer. *ecancermedicalscience* 7:343
128. Wu S, Liu S, Liu Z, Huang J, Pu X, et al. 2015. Classification of circulating tumor cells by epithelial-mesenchymal transition markers. *PLOS ONE* 10(4):e0123976
129. Ma Y, Hao S, Wang S, Zhao Y, Lim B, et al. 2015. A combinatory strategy for detection of live CTCs using microfiltration and a new telomerase-selective adenovirus. *Mol. Cancer Ther.* 14(3):835–43
130. Zamay GS, Kolovskaya OS, Zamay TN, Glazyrin YE, Krat AV, et al. 2015. Aptamers selected to postoperative lung adenocarcinoma detect circulating tumor cells in human blood. *Mol. Ther.* 23(9):1486–96
131. Desitter I, Guerrouahen BS, Benali-Furet N, Wechsler J, Jänne PA, et al. 2011. A new device for rapid isolation by size and characterization of rare circulating tumor cells. *Anticancer Res.* 31(2):427–41
132. Fernandez SV, Bingham C, Fittipaldi P, Austin L, Palazzo J, et al. 2014. TP53 mutations detected in circulating tumor cells present in the blood of metastatic triple negative breast cancer patients. *Breast Cancer Res.* 16(5):445
133. Fabbri F, Carloni S, Zoli W, Ulivi P, Gallerani G, et al. 2013. Detection and recovery of circulating colon cancer cells using a dielectrophoresis-based device: *KRAS* mutation status in pure CTCs. *Cancer Lett.* 335(1):225–31
134. Nieva J, Wendel M, Luttgen MS, Marrinucci D, Bazhenova L, et al. 2012. High-definition imaging of circulating tumor cells and associated cellular events in non–small cell lung cancer patients: a longitudinal analysis. *Phys. Biol.* 9(1):016004
135. Williams ES, Rodriguez-Bravo V, Chippada-Venkata U, De Ia Iglesia–Vicente J, Gong Y, et al. 2015. Generation of prostate cancer patient derived xenograft models from circulating tumor cells. *J. Vis. Exp.* 104:e53182
136. Khoo BL, Lee SC, Kumar P, Tan TZ, Warkiani ME, et al. 2015. Short-term expansion of breast circulating cancer cells predicts response to anti-cancer therapy. *Oncotarget* 6(17):15578–93
137. Yu M, Bardia A, Aceto N, Bersani F, Madden MW, et al. 2014. Ex vivo culture of circulating breast tumor cells for individualized testing of drug susceptibility. *Science* 345(6193):216–20
138. Hart CD, Galardi F, Pestrin M, De Luca F, Risi E, Di Leo A. 2016. Using CTCs for pharmacogenomic analysis. *Pharmacol. Res.* 106:92–100
139. Metzker ML. 2010. Sequencing technologies—the next generation. *Nat. Rev. Genet.* 11(1):31–46
140. Baslan T, Hicks J. 2014. Single cell sequencing approaches for complex biological systems. *Curr. Opin. Genet. Dev.* 26:59–65
141. Dean FB, Hosono S, Fang L, Wu X, Faruqi AF, et al. 2002. Comprehensive human genome amplification using multiple displacement amplification. *PNAS* 99(8):5261–66

142. Polzer B, Medoro G, Pasch S, Fontana F, Zorzino L, et al. 2014. Molecular profiling of single circulating tumor cells with diagnostic intention. *EMBO Mol. Med.* 6(11):1371–86
143. Heitzer E, Auer M, Gasch C, Pichler M, Ulz P, et al. 2013. Complex tumor genomes inferred from single circulating tumor cells by array-CGH and next-generation sequencing. *Cancer Res.* 73(10):2965–75
144. Dago AE, Stepansky A, Carlsson A, Luttgen M, Kendall J, et al. 2014. Rapid phenotypic and genomic change in response to therapeutic pressure in prostate cancer inferred by high content analysis of single circulating tumor cells. *PLOS ONE* 9(8):e101777
145. Gasch C, Plummer PN, Jovanovic L, McInnes LM, Wescott D, et al. 2015. Heterogeneity of miR-10b expression in circulating tumor cells. *Sci. Rep.* 5:15980
146. Steinestel J, Luedeke M, Arndt A, Schnoeller TJ, Lennerz JK, et al. 2016. Detecting predictive androgen receptor modifications in circulating prostate cancer cells. *Oncotarget*. In press
147. Riethdorf S, Müller V, Zhang L, Rau T, Loibl S, et al. 2010. Detection and HER2 expression of circulating tumor cells: prospective monitoring in breast cancer patients treated in the neoadjuvant GeparQuattro trial. *Clin. Cancer Res.* 16(9):2634–45
148. Miyamoto DT, Lee RJ, Stott SL, Ting DT, Wittner BS, et al. 2012. Androgen receptor signaling in circulating tumor cells as a marker of hormonally responsive prostate cancer. *Cancer Discov.* 2(11):995–1003
149. Kasimir-Bauer S, Bittner A-K, König L, Reiter K, Keller T, et al. 2016. Does primary neoadjuvant systemic therapy eradicate minimal residual disease? Analysis of disseminated and circulating tumor cells before and after therapy. *Breast Cancer Res.* 18(1):20
150. Fehm T, Becker S, Duerr-Stoerzer S, Sotlar K, Mueller V, et al. 2007. Determination of HER2 status using both serum HER2 levels and circulating tumor cells in patients with recurrent breast cancer whose primary tumor was HER2 negative or of unknown HER2 status. *Breast Cancer Res.* 9(5):R74
151. Cho WCS. 2014. Emerging techniques in molecular detection of circulating tumor cells. *Expert Rev. Mol. Diagn.* 14(2):131–34
152. Moreno JG, O'Hara SM, Gross S, Doyle G, Fritsche H, et al. 2001. Changes in circulating carcinoma cells in patients with metastatic prostate cancer correlate with disease status. *Urology* 58(3):386–92
153. Pailler E, Adam J, Barthélémy A, Oulhen M, Auger N, et al. 2013. Detection of circulating tumor cells harboring a unique ALK rearrangement in ALK-positive non-small-cell lung cancer. *J. Clin. Oncol.* 31(18):2273–81
154. Pailler E, Auger N, Lindsay CR, Vielh P, Islas-Morris-Hernandez A, et al. 2015. High level of chromosomal instability in circulating tumor cells of *ROS1*-rearranged non-small-cell lung cancer. *Ann. Oncol.* 26(7):1408–15
155. Antonarakis ES, Lu C, Wang H, Luber B, Nakazawa M, et al. 2014. AR-V7 and resistance to enzalutamide and abiraterone in prostate cancer. *N. Engl. J. Med.* 371(11):1028–38
156. Antonarakis ES, Lu C, Luber B, Wang H, Chen Y, et al. 2015. Androgen receptor splice variant 7 and efficacy of taxane chemotherapy in patients with metastatic castration-resistant prostate cancer. *JAMA Oncol.* 1(5):582–91
157. Wallwiener M, Riethdorf S, Hartkopf AD, Modugno C, Nees J, et al. 2014. Serial enumeration of circulating tumor cells predicts treatment response and prognosis in metastatic breast cancer: a prospective study in 393 patients. *BMC Cancer* 14:512
158. Smerage JB, Barlow WE, Hortobagyi GN, Winer EP, Leyland-Jones B, et al. 2014. Circulating tumor cells and response to chemotherapy in metastatic breast cancer: SWOG S0500. *J. Clin. Oncol.* 32(31):3483–89
159. Thalgott M, Rack B, Horn T, Heck MM, Eiber M, et al. 2015. Detection of circulating tumor cells in locally advanced high-risk prostate cancer during neoadjuvant chemotherapy and radical prostatectomy. *Anticancer Res.* 35(10):5679–85
160. Gazzaniga P, de Berardinis E, Raimondi C, Gradilone A, Busetto GM, et al. 2014. Circulating tumor cells detection has independent prognostic impact in high-risk non-muscle invasive bladder cancer. *Int. J. Cancer* 135(8):1978–82
161. He W, Wang H, Hartmann LC, Cheng J-X, Low PS. 2007. In vivo quantitation of rare circulating tumor cells by multiphoton intravital flow cytometry. *PNAS* 104(28):11760–65

162. Li J, Sharkey CC, Wun B, Liesveld J, King MR. 2016. Genetic engineering of platelets to neutralize circulating tumor cells. *J. Control. Release* 228:38–47
163. Langenbeck B. 1841. On the development of cancer in the veins. *Edinb. Med. Surg. J.* 55:251–53
164. Engell HC. 1955. Cancer cells in the circulating blood; a clinical study on the occurrence of cancer cells in the peripheral blood and in venous blood draining the tumour area at operation. *Acta Chir. Scand. Suppl.* 201:1–70
165. Seal SH. 1959. Silicone flotation: a simple quantitative method for the isolation of free-floating cancer cells from the blood. *Cancer* 12(3):590–95
166. Alexander RF, Spriggs AI. 1960. The differential diagnosis of tumour cells in circulating blood. *J. Clin. Pathol.* 13:414–24
167. Song MJ, Kornatowski G, Parsons DF, King MV. 1987. Detection and characterization of circulating rat mammary tumor cells in buffy coat and correlation with metastasis. *Cancer Investig.* 5(5):429–41
168. Schlimok G, Funke I, Holzmann B, Göttlinger G, Schmidt G, et al. 1987. Micrometastatic cancer cells in bone marrow: in vitro detection with anti-cytokeratin and in vivo labeling with anti-17-1A monoclonal antibodies. *PNAS* 84(23):8672–76
169. Hardingham JE, Kotasek D, Farmer B, Butler RN, Mi JX, et al. 1993. Immunobead-PCR: a technique for the detection of circulating tumor cells using immunomagnetic beads and the polymerase chain reaction. *Cancer Res.* 53(15):3455–58
170. Racila E, Euhus D, Weiss AJ, Rao C, McConnell J, et al. 1998. Detection and characterization of carcinoma cells in the blood. *PNAS* 95(8):4589–94
171. Vona G, Sabile A, Louha M, Sitruk V, Romana S, et al. 2000. Isolation by size of epithelial tumor cells: a new method for the immunomorphological and molecular characterization of circulating tumor cells. *Am. J. Pathol.* 156(1):57–63
172. Krivacic RT, Ladanyi A, Curry DN, Hsieh HB, Kuhn P, et al. 2004. A rare-cell detector for cancer. *PNAS* 101(29):10501–4
173. O'Hara SM, Moreno JG, Zweitzig DR, Gross S, Gomella LG, Terstappen LWMM. 2004. Multigene reverse transcription–PCR profiling of circulating tumor cells in hormone-refractory prostate cancer. *Clin. Chem.* 50(5):826–35
174. Marrinucci D, Bethel K, Lazar D, Fisher J, Huynh E, et al. 2010. Cytomorphology of circulating colorectal tumor cells: a small case series. *J. Oncol.* 2010:1–7
175. Hodgkinson CL, Morrow CJ, Li Y, Metcalf RL, Rothwell DG, et al. 2014. Tumorigenicity and genetic profiling of circulating tumor cells in small-cell lung cancer. *Nat. Med.* 20(8):897–903
176. Tulley S, Zhao Q, Dong H, Pearl ML, Chen W-T. 2016. Vita-Assay^{TM} method of enrichment and identification of circulating cancer cells/circulating tumor cells (CTCs). *Methods Mol. Biol.* 1406:107–19

II. Multiplex Protein detection on CTCs

Gerdtsson E, Pore M, Thiele J-A, Sandstrom Gerdtsson A, Malihi PD, Nevarez R, Kolatkar A, Ruiz Velasco C, Wix S, Singh M, Carlsson A, Zurita AJ, Logothetis C, Merchant AA, Hicks J and Kuhn P; 2018, **Multiplex protein detection on circulating tumor cells from liquid biopsies using imaging mass cytometry**. Converg Sci Phys Oncol 4:015002. doi: 10.1088/2057-1739/aaa013

(Journal has recently been selected for indexing in Clarivate Analytics Emerging Sources Citation Index (ESCI); listed in Web of Science as of next year).

Multiplex protein detection on circulating tumor cells from liquid biopsies using imaging mass cytometry

Convergent Science Physical Oncology

PAPER

Multiplex protein detection on circulating tumor cells from liquid biopsies using imaging mass cytometry

RECEIVED
24 September 2017

REVISED
22 November 2017

ACCEPTED FOR PUBLICATION
8 December 2017

PUBLISHED
16 January 2018

Erik Gerdtsson[1], Milind Pore[1], Jana-Aletta Thiele[2], Anna Sandström Gerdtsson[1], Paymaneh D Malihi[1], Rafael Nevarez[1], Anand Kolatkar[1], Carmen Ruiz Velasco[1], Sophia Wix[1], Mohan Singh[5], Anders Carlsson[1], Amado J Zurita[3], Christopher Logothetis[3], Akil A Merchant[5], James Hicks[1,4] and Peter Kuhn[1,4,6,7,8]

[1] Bridge@USC, USC David and Dana Dornsife College of Letters, Arts, and Sciences, 3430 S Vermont Ave, TRF 114, MC3303, Los Angeles, CA, 90089-3303, United States of America
[2] Faculty of Medicine in Pilsen, Biomedical Center, Charles University in Prague, alej Svobody 76, 323 00 Pilsen, Czechia
[3] Department of Genitourinary Medical Oncology, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Blvd, Unit 207, Houston, TX, United States of America
[4] Biological Sciences, University of Southern California, Los Angeles, CA, United States of America
[5] Division of Hematology, Department of Medicine, USC Norris Comprehensive Cancer Center, Keck School of Medicine, University of Southern California, 1450 Biggy Street, Los Angeles, CA, United States of America
[6] Aerospace and Mechanical Engineering, University of Southern California, Los Angeles, CA, United States of America
[7] Biomedical Engineering, University of Southern California, Los Angeles, CA, United States of America
[8] Medicine, University of Southern California, Los Angeles, CA, United States of America

E-mail: gerdtsso@usc.edu, pore@usc.edu, agerdtss@usc.edu, pmalihi@usc.edu, rnevarez@usc.edu, kolatkar@usc.edu, ruizvela@usc.edu, swix@usc.edu, ncarlsso@usc.edu, jameshic@usc.edu, pkuhn@usc.edu, Jana.A.Thiele@lfp.cuni.cz, azurita@mdanderson.org, clogothe@mdanderson.org, mohansin@usc.edu and akilmerc@usc.edu

Keywords: single cell analysis, imaging mass cytometry, CyTOF, protein biomarkers, HD-SCA, circulating tumor cells

Supplementary material for this article is available online

Abstract

Molecular analysis of circulating and disseminated tumor cells (CTCs/DTCs) has great potential as a means for continuous evaluation of prognosis and treatment efficacy in near-real time through minimally invasive liquid biopsies. To realize this potential, however, methods for molecular analysis of these rare cells must be developed and validated. Here, we describe the integration of imaging mass cytometry (IMC) using metal-labeled antibodies as implemented on the Fluidigm Hyperion Imaging System into the workflow of the previously established high definition single cell analysis (HD-SCA) assay for liquid biopsies, along with methods for image analysis and signal normalization. Using liquid biopsies from a metastatic prostate cancer case, we demonstrate that IMC can extend the reach of CTC characterization to include dozens of protein biomarkers, with the potential to understand a range of biological properties that could affect therapeutic response, metastasis and immune surveillance when coupled with simultaneous phenotyping of thousands of leukocytes.

Introduction

Enabling precision medicine in cancer care requires the ability to both deconvolute heterogeneity in the primary and metastatic tissues, and to characterize the liquid phase of the disease. In settings of chemotherapy, molecularly targeted and immune-system engaging treatment approaches, the high-resolution characterization of the disease along its evolutionary path at the time of the decision making is a primary challenge that can be addressed by the integration of quantitative single cell technologies within the solid and liquid biopsy workflows. The recognition of the complex interactions between the cancer and the immune system are particularly important drivers for the detailed analysis at the cellular, proteome and genome levels to both establish a treatment strategy and monitor its efficacy.

Since its commercial introduction in 2011, cytometry by time-of-flight (CyTOF) mass spectrometry using metal-labeled antibodies, has rapidly entered the basic and clinical research laboratory settings [1–3]. Its main distinction with respect to standard fluorescence-based cytometry is the capacity to assay the binding of 35 or more specific antibodies on each cell simultaneously. Bodenmiller and Günther *et al* have developed methods for coupling laser ablation

with CyTOF technology, enabling multiplexed image-based proteomic analysis of formalin-fixed paraffin-embedded (FFPE) tissue sections or cultured cells mounted on glass slides [4, 5]. The technology is now commercially available from Fluidigm Corporation (South San Francisco) as the Hyperion Imaging System for imaging mass cytometry, or IMC. In the IMC process, the target tissue or cell preparation is treated with a cocktail of antibodies, each labeled with a specific rare earth isotope. Regions of interest (ROI) on the slides are scanned *in situ* with a highly focused, pulsed laser, such that each pulse vaporizes a 1 μm^2 bloc of the sample and the resulting ions are introduced into the inductively coupled plasma time-of-flight mass spectrometer (ICP-TOF-MS) with helium as a carrier gas. The ion counts for each pulse can then be assembled into a protein expression image with a resolution of 1 μm^2 across the ROI allowing for limited characterization of sub-cellular localization.

In this report, we describe the integration of imaging mass cytometry (IMC) using the Hyperion instrument into the previously validated high definition single cell analysis (HD-SCA) workflow [6], with the goal to incorporate morphology, proteomics and genomics of rare single cells in a single streamlined process. The HD-SCA method was designed to identify and characterize ultra-rare (<0.0001%) cancer cells in liquid biopsies from blood and bone marrow aspirates, or high-complexity samples such as tissue 'touch preparations' on glass slides. In contrast to other CTC technologies, the HD-SCA workflow uses a direct analysis, 'no cell left behind' strategy in which the entire population of nucleated cells from a blood draw is spread on glass slides at 3 million cells per slide, stained, and imaged *in situ*. Fresh samples must be fractionated within 48 h following blood collection, which enables simple shipping from clinical sites. Candidate tumor cells are identified among the millions of leukocytes using immunofluorescent markers. In addition to the morphometric information obtained by imaging, the candidate circulating tumor cells (CTC) from blood or disseminated tumor cells (DTC) from bone marrow, or other cells of interest can subsequently be relocated and physically isolated for single cell genomic analysis [7–9]. The data derived from the HD-SCA workflow has been demonstrated to generate clinically actionable information [10]. Further, due to their minimally invasive nature, liquid biopsies can be repeated frequently during treatment to assess response and provide early indications of changes that could lead to disease progression. Despite its high sensitivity, the phenotyping of cells detected by HD-SCA has been largely limited to morphometry and four fluorescent markers. Targeted single cell proteomic analysis can complement genomic and whole plasma characterization to assign functional relevance to the molecular observations as part of an integrated signature. The longitudinal characterization of the disease using cancer cells from blood and bone marrow liquid biopsies provides insight into the evolution of the tumor, and the significance of sub-populations of circulating tumor cells present in specific environments.

Here, we present experimental procedures along with data analysis methods, specifically adapted for the HD-SCA workflow to incorporate multiplex targeted proteomic analysis on intact single cells using the Hyperion instrument. The functionality and performance of the workflow is demonstrated, including methods for validating metal conjugated antibodies for use specifically in the combined HD-SCA/IMC workflow. In addition, the robustness of the method allows for immunophenotyping of the leukocyte populations that are simultaneously imaged along with the CTCs, which is demonstrated here by gating the multi-parametric data from the HD-SCA/IMC workflow in a manner similar to cytometric sorting methods, yielding well separated sub-populations. Finally, we present combined proteo-genomic results relating the HD-SCA/IMC to the HD-SCA/genomics from the analysis of liquid biopsies from a metastatic prostate cancer patient, showing the distribution and subcellular location of characteristic prostate cancer protein markers in a clonally related population of CTCs and DTCs.

Methods

HD-SCA specimen preparation

Whole blood samples from primary and metastatic cancer patients were collected in cell-free DNA blood collection tubes (Streck, Omaha, USA) and shipped to the Kuhn laboratory at USC in temperature stabilized shippers (Standard 71, LLC, Los Angeles). Samples were processed within 24 h of collection using the HD-SCA protocol [6] and archived at −80 °C. The sample MDA-42109 was collected from a patient with metastatic prostate cancer in 2013 with informed consent and under IRB approval as previously described [8]. Normal control donor blood samples were obtained from The Scripps Research Institute, La Jolla, California, USA. All specimen collections were performed under IRB approval.

Cell lines spiked in blood

Cell lines were obtained from the American Type of Culture Collection, (ATCC, Manassas, VA, USA). Prostate cancer line LNCaP was cultured in RPMI-1640 medium and breast cancer line MDA-MB-231 was cultured in DMEM/F12 medium (Corning, Manassas, VA, USA). Cells were harvested using Versene solution (Gibco) and washed in PBS. For validation of the metal-labeled antibodies in the combined workflow, cultured cells were added at a ratio 1:10 000 to whole blood from normal blood donors and processed according to the standard HD-SCA protocol [6], including storage at −80 °C for a minimum of 24 h.

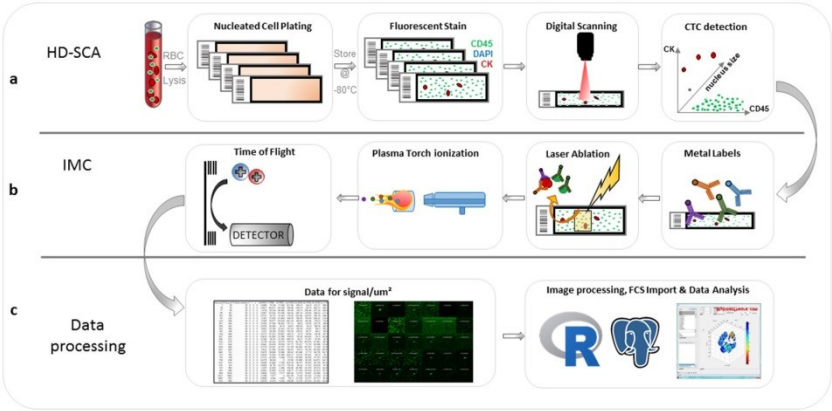

Figure 1. Schematic workflow for protein profiling of rare cells identified with the HD-SCA workflow. (a) Sample processing and fluorescent imaging; (b) IMC staining and data acquisition; (c) data processing.

HD-SCA fluorescent labelling, imaging and cell identification

To complete the HD-SCA procedure, slides were thawed, stained with DAPI (nuclear stain) and fluorescent antibodies against CD45 (leukocyte marker) and pan-cytokeratin (CK, epithelial marker), and imaged in three colors using a high-speed automated microscope scanner. As part of the normal HD-SCA image analysis, the candidate CK positive cells were identified and coordinates registered. High resolution 40× images of candidate cells were captured using a Nikon 80i microscope. Prior to staining with the metal-labeled antibody cocktail for IMC analysis, the slides were stored at 4 °C for several days to parallel normal operation of the HD-SCA protocol. Finally, cover slips were removed in PBS with minimal disruption of the cell layer. Slides were then washed with PBS to dissolve the imaging media in preparation for the metal-labeled antibody incubation. For the HD-SCA/Genomics workflow results presented in figure 3, samples were stained with 4-colors, adding AR to the DAPI, CD45 and CK panel described above.

MaxPar metal-labeled antibody staining

Metal-labeled antibodies were provided by Fluidigm, either from the standard CyTOF catalog (http://maxpar.fluidigm.com/product-catalog-metal.php) or as custom conjugates (supplementary table 1 (stacks.iop.org/CSPO/2/015002/mmedia)). Metal-labeled antibody cocktails were prepared in 0.1% Tween-20, 1% BSA in PBS as per the dilution scheme presented in supplementary table 1. All samples, prepared as described above, were first blocked with 1% BSA and 0.2 mg ml^{-1} mouse IgG Fc fragment (Thermo Scientific) in PBS for 30 min and then incubated with antibody cocktail for 1.5 h at RT, followed by washing with PBS and staining with DNA intercalator Ir-193 (Fluidigm) and cell membrane counter-stain In-115 (Fluidigm) for 30 min. Slides were again washed with PBS and rinsed with ddH20 for 5 s and dried overnight at room temperature prior to IMC analysis.

Statistical considerations and data analysis

Cell boundaries were determined by the segmentation method described in the results section, using the Bioconductor package EBImage in R [11]. The standard deviation over the mean (SDOM) per cell for each channel was defined as the difference between the average ion count for the cell and the mean signal for all the cells divided by the standard deviation of all cells within each ROI (400 × 400 μm). The noise level, used as a denominator for the S/N ratio was approximated as 1.5 × IQR (inter-quartile range) outside the upper quartile of all cells. When applicable, the limit of detection (LOD) for each marker was set as to mean + 3.3 × standard deviation (one sided 95% ∗ 2) or S/N = 3 [12]. For figure 5, the dataset was exported in a FCS format using the FlowCore package in R and the gating viSNE plots were generated with the CYT graphical software package [13] in MATLAB.

Results

The HD-SCA/IMC integrated workflow

Both the HD-SCA and the IMC were designed to operate on standard glass microscope slides in keeping with general procedures in the clinical pathology laboratory setting. The common slide format facilitated the integration of the two technologies into a single workflow viable for both basic and clinical research applications. The workflow, shown in figure 1(a), begins with processing the blood sample and spreading all nucleated cells on slides, followed by 3- or 4-color immunofluorescent staining and high-speed digital scanning to identify and image cells of

interest. The coordinates of each candidate cell are recorded and re-imaged at 40× for morphometric measurements, as previously reported [6, 8]. In preparation for the IMC (figure 1(b)), slides are re-incubated with MaxPar™ metal-labeled antibodies. The slide is placed in the IMC instrument and candidate cells are relocated using a previously determined offset between coordinates from the 40× imaging microscope and the IMC stage. A 400 × 400 μm ROI around the cell of interest undergoes laser ablation aerosolizing a 1 μm^2 area/pulse (200 Hz), followed by ionization and quantification in the CyTOF Helios instrument. Ion mass data is collected for each pulse and processed to render images for each individual channel at 1 μm resolution, where the intensity of each pixel corresponds to the ion count value. The images are analyzed using an in-house developed software application in R. Raw data and processed results are stored in a PostgreSQL database enabling export of datasets and integration with HD-SCA assay parameters (figure 1(c)).

Although the basic formats of the two assays are similar, adapting the IMC technology to the HD-SCA workflow required the development and validation of staining and analysis methods consistent with the previously validated HD-SCA protocols. The antibody preparation and staining steps were initially developed and optimized using cell lines spiked into normal donor blood samples (Methods).

For this study, a panel of metal-labeled antibodies was designed to include targets relevant to prostate cancer. The panel included PSA, PSMA and two epitopes for AR, along with epithelial markers EpCAM, E-cadherin, cancer stem cell marker CD44, EMT marker B-catenin, as well as a subset of markers for leukocyte characterization [14–16]. HLA-DR and CD14 were selected for monocytes identification along with CD3, CD8, CD4, CD45RA and CD38 to further characterize the T cell population [17]. CD66 was selected to identify mainly the neutrophil population, however, CD66 or CEA (carcinoembryonic antigen) is also expressed in in tumor cells and is known to have a tumor suppressive role in prostate cancer [18]. All catalog antibodies had been previously validated by Fluidigm for use in the CyTOF/Helios systems, but had not been validated in the IMC modality nor tested in the context of the HD-SCA workflow.

Each metal-labeled antibody was first tested for specificity and sensitivity against the prostate cancer line LNCaP, and metastatic breast cancer line MDA-MB231 (Materials and Methods and supplemental figure S1), using the leukocyte populations surrounding the candidate cells as an additional control. The unique challenge of determining the presence of the various antigens on the single candidate cell within each ROI in comparison to flow-based methods required the development of routines for quantification and scoring as described below.

Cell segmentation and statistical analysis

As the pulsed laser sequentially ablates subcellular areas across the candidate cell along with all of the ~300 leukocytes present in the defined ROI, an ion count value is reported for each isotope at each of the 160 000 acquired pulses. The resulting data matrix is stored in a text file from which false color 1 μm resolution images are generated from each array of ion count values. The range of ion count values is internally normalized within each image, usually so that the brightest pixel represents the 98th percentile of the cumulative ion count signal.

Figure 2 shows an example of the images obtained from the IMC analysis of the two spiked cell lines. In the top row, an aggregate of LNCaP cells is seen in the center of each frame. A single LNCaP cell is also captured within the ROI. The different localized signal for PSA (membrane, figure 2(e)), E-cadherin (cell junctions, figure 2(f)), and androgen receptor (AR) (translocated to nuclei upon activation, figure 2(g)), can clearly be seen. For segmentation, the DNA intercalator data (figure 2(d)) is used to identify the objects in each frame. Adaptive thresholding [11] is utilized to generate a binary image, and then applying a watershed algorithm [19] to separate objects that are touching. Finally, the membrane counterstain is used for Voronoi propagation [20]to define the cell boundaries of each object with the primary mask acting as seeds (figure 2(c)). The mask is used to quantify the signals associated with each cell for each channel, by applying the mask on the raw data array of ion count values, rather than the normalized image. The ion count signals within the boundaries of each segmented cell in the ablation field are used to generate biaxial plots showing the average signal per cell for each marker relative to a chosen leukocyte marker, typically CD45 (figures 2(m)–(o)). This quantitation leads to a heuristic scoring system for each marker as described below.

Design of a scoring system for marker classification

In practice, the quantitation of different markers is hampered by (i) subjectivity in adjusting the gain for normalization of images, (ii) a wide dynamic range of protein expression, (iii) inherent differences in detector sensitivity for the metal ions [21], (iv) experimental variation due to instrument fluctuation [22] and differences between staining batches etc, and (v) the difficulty of separating low abundant protein signals from background noise. As alluded to above, these challenges exacerbated in the setting of evaluating rare single cells, as compared to the common application of cytometry where larger populations are characterized.

Classification based on ion count alone was deemed unfeasible due to vast differences in dynamic range for each antibody-analyte pair, impeding normalization of the signals and their interpretation in heat maps. Thus, to establish an objective, semi-automated means for evaluating signals, a four level scoring system was created, with 0 being below the

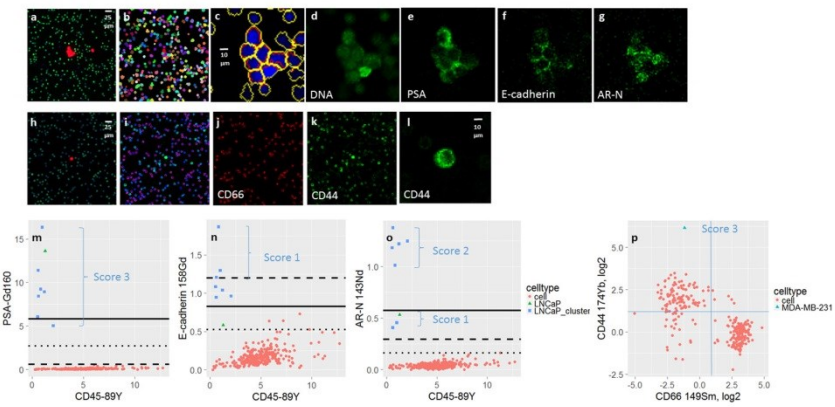

Figure 2. Illustrations of segmentation and scoring of individual cells and aggregates. (a) HD-SCA composite image of LNCaP cells spiked in normal blood containing a cluster of cells in the center of the frame along with a single cell to the right. (b) Colored segmented cells in the IMC ROI. (c) Zoomed in LNCaP cluster of cells segmented using the DNA-Intercalator signal (blue) to outline the nuclei (yellow) as primary masks acting as seeds for propagation onto the membrane image to identify the outer cell boundaries (red). (d)–(g) Images showing the signal for individual IMC channels as labeled. (h) HD-SCA composite image of a single MDA-MB-231 cell spiked in normal blood. (i) Composite image, showing CD66 (red), CD44 (green) and DNA (blue), demonstrating the mutually exclusive populations of leukocytes (j) and (k). (l) Enlarged image of the MDA-MB-231 cell stained with CD44. (m)–(o) Biaxial plots of the segmented cells from ROI in figure (b) with channels specified on the *y*-axis and CD45 on *x*-axis. The LNCaP cells are highlighted as green triangles (part of the cluster) or blue squares (single cell), with the associated score. The lines represents the estimated LOD based on either S/N (dashed), SDOM for all cells (solid) or SDOM based on the leukocytes only (dotted). (p) Biaxial plots of CD66 versus CD44 on a log2 scale of the cells present in the ROI in I along with gating lines in blue. The MDA-MB-231 cell is shown as a blue triangle, scored as a 3.

LOD, 1 exceeding LOD, and levels 2 to 3 set to discriminate the highest signal levels. The scoring system is based both on SDOM, and signal to noise (S/N) of the ion count value for the candidate cell to the surrounding leukocyte population, providing a relative measurement that is less sensitive to experimental variation than ion count alone. The LOD for each marker was set as equal to S/N = 3 or SDOM = 3.3 [17]. Cells exceeding the LOD were scored as 1+, signals with S/N of 7-20 or SDOM > 6 were scored 2+, and finally score 3+ was assigned to signals of S/N > 20 or SDOM > 12. In figures 2(m)–(p), the scores are marked in the plots along with the estimated LOD based on either S/N = 3 (dashed lines) or SDOM = 3.3.

Some of the markers tested, including CD44, CD66, and vimentin, are also present on subpopulations of leukocytes. As seen in figures 2(i)–(k), most of the cells in the ROI express either CD44 or CD66, comprising two mutually exclusive subpopulations. In such cases, manual or automated gating can be readily applied to score the cell of interest, as demonstrated for a spiked MDA-MB-231 cell shown in figure 2(p). In this case, the CD44 signal was above the gated CD44+ leukocyte population, and was given a score of 3. Additional scoring of the different markers for LNCaP and MDA-MB-231 are presented in supplemental figure 1.

Analysis of a liquid biopsy sample from a prostate cancer patient

To test the integration of IMC with the HD-SCA workflow and to demonstrate the added information provided by IMC over 4-color immunofluorescence, the combined HD-SCA/IMC workflow was performed on matched blood and bone marrow aspirate samples from a patient with metastatic castrate resistant prostate cancer (MDA-42109) collected as part of a larger study [8]. In the initial HD-SCA assay of MDA-42109 conducted in 2013, 9 CTCs ml^{-1} from blood and 207 DTCs ml^{-1} in bone marrow were identified, with 48% of the CTCs and 78% of the DTCs also scoring positive for AR expression [23]. High resolution (40×) images had been collected for all the CTCs and a subset of DTCs to verify the AR positivity and subcellular localization prior to isolating the cells for whole genome amplification using the HD-SCA genomics workflow [7]. The results showed that 41/42 of the successfully amplified cells yielded highly rearranged and closely related copy number profiles with common elements typical of metastatic prostate cancer, indicating a single cancer lineage with four distinct sub clones as shown in the heat map and representative profiles (figure 3(a)). The high degree of clonality in the CNV profiled cells provided assurance that the candidate CTCs and DTCs selected for subsequent IMC analysis were highly likely to be *bona fide* cancer cells from the same lineage. However, we note that despite sharing the same genotype, individual cells exhibit significant phenotypic variability in AR expression and subcellular location, consistent with previous results [7]. In most cases, amplification of the AR gene on chromosome X corresponds to a strong nuclear AR signal in the corresponding fluorescent

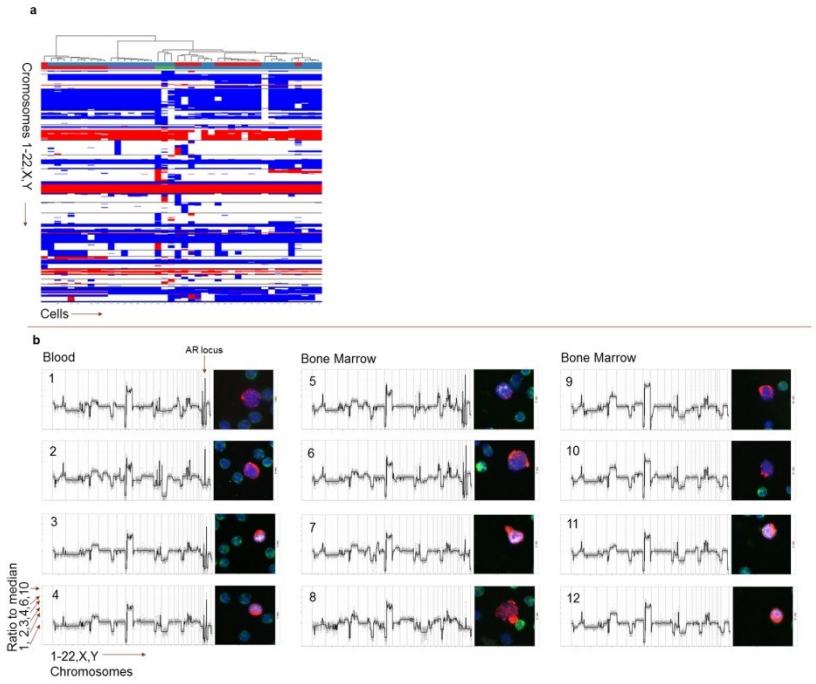

Figure 3. Genomic profiles of cells found in the blood and bone marrow from the prostate cancer patient. (a) Genomic heat map displays the CNV profiles (chromosomal deletions in blue and amplification in red) for the CTC/DTC analyzed from the prostate cancer patient. (b) Representative CNV profiles of a CTC (left) and DTC (center, right) including 40× immunofluorescent composite images (DAPI, blue; CK, red; CD45, green; and AR white). The AR loci on X chromosome is highlighted ((b), panel 1).

image (figure 3(b), panels 3, 4, 5 and 7). In other cases, the gene encoding AR was amplified even though the protein could not be detected by fluorescence (figure 3(b), panels 1, 2 and 6). Notably, there is also a genomic subclone comprising cases where AR amplification is not observed in the CNV profile (figure 3(b), panels 9–12) but AR is still detected in some of the fluorescent images (figure 3(b), panels 11 and 12). Morphological traits, such as cell and nuclear size and CK staining also varied among individual cells as is evident from the fluorescent images.

To evaluate the performance of the combined HD-SCA/IMC analysis in relation to the results from the HD-SCA fluorescent assay alone, we retrieved one unstained slide each from blood and bone marrow of MD42109 from the −80 °C archive where they had been stored for three years. The slides were stained and scanned using the standard HD-SCA 3-color protocol (DAPI, CD45 and pan-CK) to identify candidate cancer cells (figure 1). A total of 21 candidate DTCs and three CTCs were identified and their positions located on the respective slides. The stained slides were then stained with metal-labeled antibodies in preparation for imaging on the Hyperion instrument. The combined results of the HD-SCA immunofluorescence and IMC analysis of these cells are presented in figures 4 and 5.

Figures 4(a) and (b) show a selection of images generated in the IMC analysis of two representative cells, DTC #1271 from the bone marrow aspirate and CTC #1290 from blood, along with separate images of the markers being tested in the two modalities. The images from the fluorescence analysis of the representative cells (top panel) show that both cells were strongly CK positive and CD45 negative. The CD45 expression in leukocytes seen in the full ROI composite image generated by the IMC, show a high degree of correlation of CD45 expression ranging from low to high with the same cells in the corresponding ROI captured in the HD-SCA fluorescence assay image (figure 4(a), supplementary data, figure 1). The composite fluorescence images (figure 4(c)), demonstrate a phenotypic heterogeneity of both the morphology (size) and the variation in the CK signal among the cancer cells.

The heat map (figure 4(d)) provides a summary of the IMC signals in each channel for all 24 cells analyzed. The signal intensity for CK 8/18 varies even among the positive cells, and some of the cells that were positive for CK in the HD-SCA analysis were below the LOD in the IMC for CK 8/18, likely due to the fact that more CK epitopes are measured with the pan-CK antibody cocktail used in HD-SCA staining.

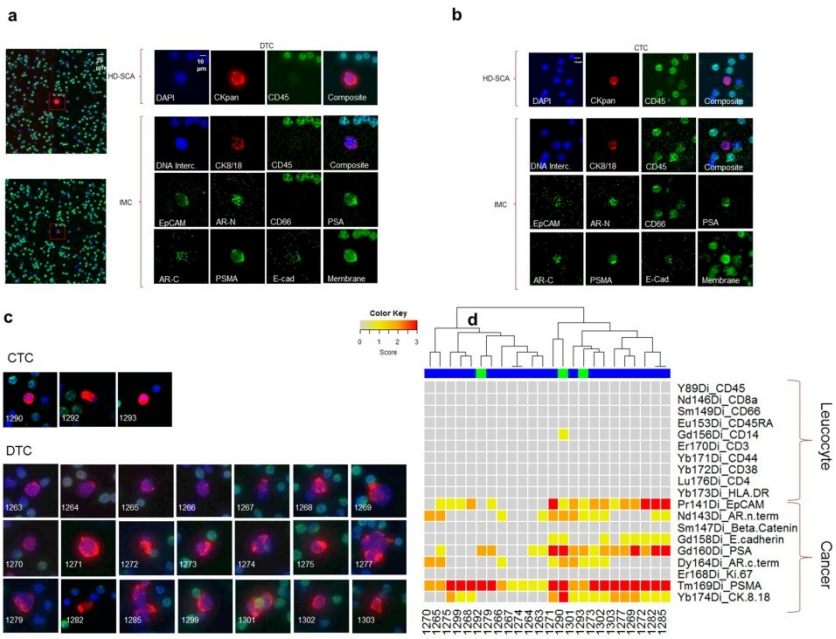

Figure 4. Combined IMC and HD-SCA characterization of circulating tumor cells from prostate cancer blood and bone marrow. (a) DTC (#1271) from a prostate cancer patient centered in the full frame ROI (400 × 400) together with cropped/zoomed in fluorescent images (top panel) and the IMC rendered images (bottom panels) as comparison for channel DAPI/Intercalator (blue), CK (red), CD45 (green) and as a 3 color composite image. The IMC generated results for EpCAM, AR-N, CD66, PSA, PSMA, E-cadherin and a lipid membrane counterstain are presented in green below. (b) Images of the CTC (#1290) with a similar expression profile from the same patient as in A. (c) 40× fluorescent images DAPI (blue), CD45 (green), and CK (red) of all the cells found and analyzed in the blood and bone marrow sample taken with optimal exposure for each cell. (d) Heat map of CTCs (green banner) and DTCs (blue banner) analyzed from the same patient colored according to the four step scoring scale; below detection (gray) detectable (yellow), strong (orange) and very strong (red) depending on the detected protein.

Fifty percent (12/24) of the cells showed AR positivity in the IMC, in agreement with the parallel HD-SCA assay. Of the AR N-terminal positive cells, the majority were also positive for the C-terminal specific antibody, indicating that the full-length receptor is present in these cells. The rendered images from the IMC show a clear nuclear localization of AR, which follows upon activation of the receptor [24] (figures 4(a) and (b)). A majority of the cells (18/24) were positive for the epithelial marker EpCAM and 6 cells were also positive for E-cadherin, including the cell clusters 1277 and 1285 (figures 4(c) and (d)). Importantly, all of the cells were PSMA positive and a majority were also PSA positive, providing independent evidence for the prostate origin of the cells (figures 4(a), (b) and (d)). Thus, the IMC analysis confirmed and complimented the fluorescent HD-SCA and genomic signatures by providing additional markers for phenotypic profiling. Further, these results demonstrate the feasibility of using archived samples for IMC analysis, allowing us to revisit and further characterize long stored and previously analyzed samples in the HD-SCA biobank.

Characterization of leukocyte populations

The HD-SCA assay is unique among the current CTC isolation methods in that it also assays and displays the full complement of leukocytes along with candidate CTC, enabling the analysis of a significant number of immune cells for multiple markers while analyzing CTCs. To demonstrate the feasibility of such an application, the surrounding leukocytes from the prostate cancer samples presented above were characterized by combining data aggregated from multiple ROIs on the slide into one dataset for each of the blood and bone marrow samples. To establish a reproducible analytical workflow, the first step was to remove all poorly segmented cells by size selection. Next, by gating the data (2089 cells) based on the average ion counts for each leukocyte on biaxial plots, the ratio compared to all nucleated cells of neutrophils ($CD66^{+}/CD44^{low}$, 67%) monocytes ($HLA\text{-}DR^{+}/CD14^{+}$, 7%) and T-cells ($CD3^{+}/HLA\text{-}DR^{-}$, 9%) could be determined (figure 5(a)). The ratio correlated well with the manual differential count performed by a trained clinical lab technician (69% neutrophils, 14% lymphocytes, 8% eosinophils, 8%

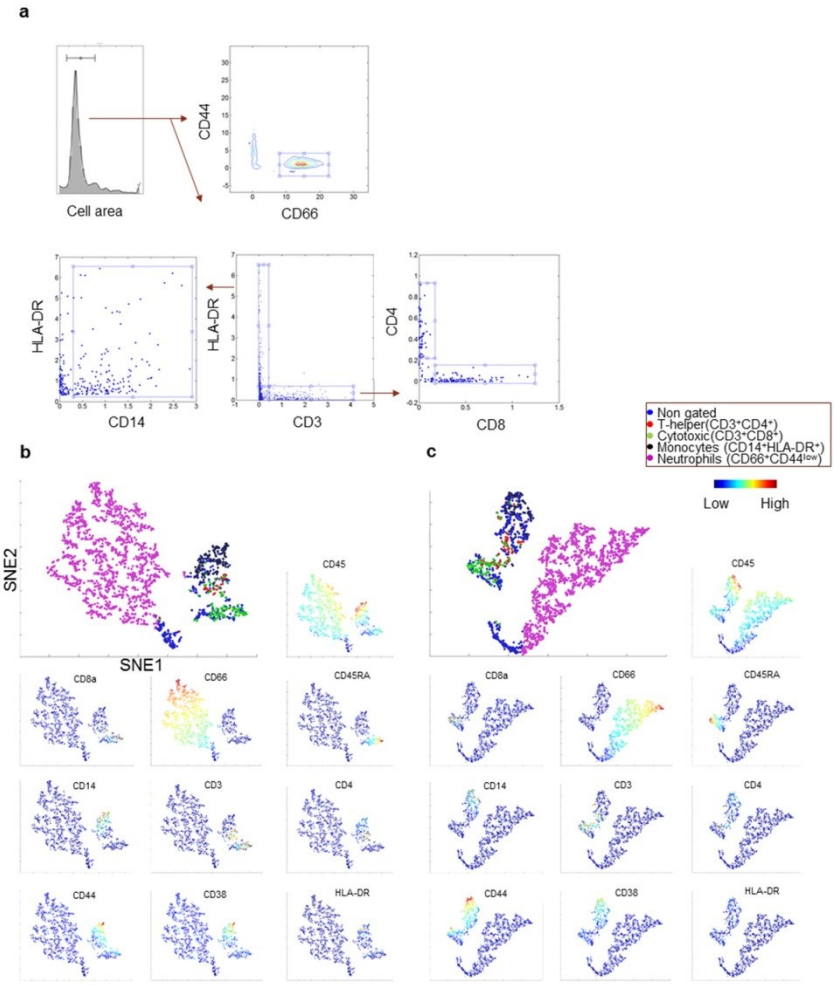

Figure 5. Immune cell phenotyping. Identification of subpopulations of leukocytes within the ROIs analyzed for sample MDA-42109 showing (a) conventional biaxial gating of cells (after initial size exclusion of non-single cells) to identify neutrophils ($CD66^+$), cytotoxic T-cells ($CD3^+$, $CD8^+$), T-helper cells ($CD3^+$, $CD4^+$) and monocyte populations ($HLA\text{-}DR^+$, $CD14^+$). (b) viSNE plots (SNE1 versus SNE2) of around 2000 cells captured with IMC in blood and (C) bone marrow, respectively. The larger image shows the same populations as gated above, overlaid by color. The smaller images show the individual markers with the expression represented as a gradient from low (blue) to high (red).

monocytes, 1% basophils). Furthermore, the T-helper $CD4^+$ cells (2%) and cytotoxic $CD8^+$ (5%) were also identified within the T-cell population. Amongst each of the T-helper and cytotoxic T-cell populations, we also detected small populations of activated ($CD38^+$, <1%) and naïve/effector cells ($CD45RA^+/CD4^+$, 1% and $CD45RA^+/CD8^+$, 3.8%). In figure 5(b), the signal intensity for different populations are shown overlaid in a 2D-viSNE map [13]. Similar data analysis for the bone marrow sample (2267 cells, figure 5(c)) demonstrated slightly lower ratio of neutrophils (55%), monocytes (4%) and lymphocytes (10%) compared to the blood. Thus, despite lacking markers for precursor cells in this particular analysis, the current results still indicate higher levels of immature cells present in the bone marrow compared to the blood, as expected [25].

In summary, we have developed an integrated approach of CTC and DTC characterization based on the HD-SCA workflow with the subsequent downstream multiplex proteomic analysis using IMC. The method was demonstrated using CTCs, DTCs as well as surrounding leukocytes in samples from a bone metastatic prostate cancer patient, showing added

value to the data generated by the HD-SCA assay alone. Results were compared to genomic profiling of matching cells from the same samples, showing a higher degree of heterogeneity on a protein level than on the genomic level.

Discussion

The results presented here provide proof of concept for the application of the combined HD-SCA/IMC workflow for targeted proteomics analysis on rare cancer cells and cell aggregates, as well as associated leukocyte populations, using blood and bone marrow aspirate from a metastatic prostate cancer patient. To achieve these results, we have coupled IMC, using the Fluidigm Hyperion instrument, with the existing HD-SCA method for CTC/DTC identification, morphometrics and single cell genomics. Significantly, this additional downstream high-content analysis was integrated without the need to modify any of the previously validated upstream processes of the HD-SCA. This continuous workflow enables a strategy of deep proteo-genomic analysis applicable for translational science studies within clinical trials. While the technology is currently applied as a discovery tool for markers or new combinations of markers, it has the potential for clinical validity either as a broader pathology platform, or during the development phase of a specific test. Typically, after discovery of a potential biomarker using such a high dimensional technology, a less complex, more focused assay based on a limited number of reagents and suitable for routine clinical implementation would likely be created, as demonstrated by Scher and colleagues [10].

The opportunity to utilize targeted proteomics for research on ultra-rare cells such as CTCs from liquid biopsies lies in the high level of multiplexing available to create 'smart panels' with a wide, purposeful spectrum of targets such as tissue specific markers and known markers of therapeutic importance, but also markers of discovery, such as mechanism of immune surveillance and escape [26] and markers for tumor cell states, including, dormancy, EMT, vascular mimicry and stemness. In practice, unique panels would be constructed for specific clinical studies. In this demonstration study, for example, we included markers for characterization of the surrounding leukocyte population. Patient specific immune signatures in addition to the cancer cells could potentially provide insight for classification and monitoring of cancer for precision medicine purposes [27, 28], and the possibility to do so simultaneously is a unique feature of the HD-SCA assay compared to other CTC technologies. The current panel contained four prostate specific markers, six epithelial and cell signaling or cancer-associated markers, and nine common leukocyte markers, plus a direct conjugated DNA intercalator and membrane counterstain. This combination of markers allowed us to both confirm the tissue of origin for the candidate prostate cancer cells and identify specific populations in the surrounding leukocytes, and further, to test the expression of additional proteins in the candidate cells.

In classifying rare cells it is important to establish normalization and quantification methods to enable cell to cell comparisons as well as comparisons between experiments. This is especially true for those proteins where expression approaches the LOD. Each individual antibody assay has to be evaluated in order to estimate the detection limit and linear range. Quantitative scoring of protein expression in individual candidate cells is an ongoing process. The scoring method presented in this study is rooted in that developed for the fluorescent measurements in the HD-SCA, wherein the signals from the surrounding white blood cells are used to establish a negative control baseline for most epithelial tissue markers and a means for normalizing signals from candidate cells. The SDOM measure as applied in the conventional HD-SCA assay is a straightforward means for establishing the background level based on the surrounding leukocytes. SDOM is however sensitive to outliers and is poorly suited for skewed populations, in particular for signals close to zero. SDOM is also limited in the case of CTC clusters, where multiple cells of interest contribute to the overall mean and standard deviation, as illustrated in figures 2(m)–(o) where the LOD (mean + 3.3σ) is estimated by excluding the seven cluster cells from the overall calculation (dotted lines) in comparison to estimate based on all cells (solid lines). The S/N measure based on the inter-quartile range was introduced because it is less sensitive to outliers, but on the other hand it gives only a rough estimate based on the background signal. Hence, both measures are currently taken into account to provide an objective score.

It should not be overlooked that each type of preparation, whether fresh frozen tissue, FFPE sections or the HD-SCA, requires separate validation of the antibody repertoire [29, 30]. In the process of integrating the two platforms, we noted that up to half of the antibody conjugates previously validated for suspension CyTOF did not work well or at all on the glass substrate HD-SCA preparations. We have therefore developed a standardized assay development process for the HD-SCA/IMC workflow using cell lines spiked into normal blood samples. Optimization of sensitivity and specificity for some problematic markers often required iterative testing of multiple antibody clones and metal-conjugate combinations. Expanding the repertoire of validated conjugates in the HD-SCA/IMC setting is a work in progress.

Rare cells from liquid biopsies show great potential for assessing the effectiveness of cancer therapy and detecting resistance or progression in near-real time through a minimally invasive procedure that can be repeated on a time scale of days or weeks, rather than months or years. New technologies such as cell free

DNA and single cell sequencing add volumes of new information to what was once limited to enumerating cancer cells in the blood. The HD-SCA workflow is a direct analysis platform that provides the opportunity for a complete characterization of the blood sample without enrichment and that accommodates multiple assay modes including both morphometric and molecular level analysis. The integration of the HD-SCA workflow with single cell targeted protein analysis creates a true multiplex technology for liquid biopsies that is accessible for both basic and translational clinical studies.

Conclusion

This study provides proof of concept that IMC analysis adds another dimension to the established HD-SCA workflow, enabling multiplex proteomic profiling in addition to morphometric and genomic characterization of ultra-rare circulating tumor cells. Importantly, this study provides evidence that samples stored for several years can be revisited and analyzed *de novo* as new protein targets are identified. With a multiplexing level of (currently) 40 proteins, highly informative biomarker panels can be readily tailored for the scientific question at hand, which gives the HD-SCA/IMC technology immense potential for deriving predictive and prognostic biomarkers from the liquid phase of cancer.

Acknowledgments

This work was based wholly or partially on research supported by Breast Cancer Research Foundation 004698-00003; National Cancer Institute and Leidos Biomedical Research, Inc HHSN261200800001E; Prostate Cancer Foundation 16CHAL04; Vassiliadis Research Fellowship, Polak Research Fellowship, the Vicky Joseph Research Fellowship, and the Charles University Research Fund (Progres Q39). USC is an early access partner of Fluidigm for the evaluation of the Hyperion System and Fluidigm provided both expertise and reagents to support this work. The content is solely the responsibility of the authors and does not necessarily represent the official views of these funding agencies and foundations. Competing Financial Interests are PK: Advisor to Epic Sciences, royalty recipient from Epic Sciences, shareholder at Epic Sciences; JBH: On the Clinical Advisory Boards of Epic Sciences, Inc., La Jolla, CA and CelMatix, Inc. of NY, NY.

ORCID iDs

Anna Sandström Gerdtsson https://orcid.org/0000-0003-1932-0365
Carmen Ruiz Velasco https://orcid.org/0000-0003-2370-3000
Peter Kuhn https://orcid.org/0000-0003-2629-4505

References

[1] Baca Q, Cosma A, Nolan G and Gaudilliere B 2017 The road ahead: implementing mass cytometry in clinical studies, one cell at a time *Cytometry* B **92** 10–1
[2] Spitzer M H and Nolan G P 2016 Mass cytometry: single cells, many features *Cell* **165** 780–91
[3] Bandura D R, Baranov V I, Ornatsky O I, Antonov A, Kinach R, Lou X, Pavlov S, Vorobiev S, Dick J E and Tanner S D 2009 Mass cytometry: technique for real time single cell multitarget immunoassay based on inductively coupled plasma time-of-flight mass spectrometry *Anal. Chem.* **81** 6813–22
[4] Wang H A, Grolimund D, Giesen C, Borca C N, Shaw-Stewart J R, Bodenmiller B and Günther D 2013 Fast chemical imaging at high spatial resolution by laser ablation inductively coupled plasma mass spectrometry *Anal. Chem.* **85** 10107–16
[5] Giesen C *et al* 2014 Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry *Nat. Methods* **11** 417–22
[6] Marrinucci D *et al* 2012 Fluid biopsy in patients with metastatic prostate, pancreatic and breast cancers *Phys. Biol.* **9** 016003
[7] Dago A E *et al* 2014 Rapid phenotypic and genomic change in response to therapeutic pressure in prostate cancer inferred by high content analysis of single circulating tumor cells *PloS One* **9** e101777
[8] Carlsson A, Kuhn P, Luttgen M S, Dizon K K, Troncoso P, Corn P G, Kolatkar A, Hicks J B, Logothetis C J and Zurita A J 2017 Paired high-content analysis of prostate cancer cells in bone marrow and blood characterizes increased androgen receptor expression in tumor cell clusters *Clin. Cancer Res.* **23** 1722–32
[9] Malihi P D *et al* 2017 Clonal diversity revealed by morphoproteomic and copy number profiles of single prostate cancer cells at diagnosis *Converg. Sci. Phys. Oncol.* **4** 015003
[10] Scher H I *et al* 2016 Association of AR-V7 on circulating tumor cells as a treatment-specific biomarker with outcomes and survival in castration-resistant prostate cancer *JAMA Oncol.* **2** 1441–9
[11] Pau G, Fuchs F, Sklyar O, Boutros M and Huber W 2010 EBImage—an R package for image processing with applications to cellular phenotypes *Bioinformatics* **26** 979–81
[12] International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) 1997 Validation of Analytical Procedures: Methodology Q2B *Fed. Reg.* **62** 27463–7 (www.ich.org/products/guidelines/quality/quality-single/article/validation-of-analytical-procedures-text-and-methodology.html)
[13] Amir E A D, Davis K L, Tadmor M D, Simonds E F, Levine J H, Bendall S C, Shenfeld D K, Krishnaswamy S, Nolan G P and Pe'er D 2013 viSNE enables visualization of high dimensional single-cell data and reveals phenotypic heterogeneity of leukemia *Nat. Biotechnol.* **31** 545–52
[14] Ghosh A and Heston W D 2004 Tumor target prostate specific membrane antigen (PSMA) and its regulation in prostate cancer *J. Cell. Biochem.* **91** 528–39
[15] Jiang Y G, Luo Y, He D L, Li X, Zhang L L, Peng T, Li M C and Lin Y H 2007 Role of Wnt/beta-catenin signaling pathway in epithelial-mesenchymal transition of human prostate cancer induced by hypoxia-inducible factor-1alpha *Int. J. Urol.* **14** 1034–9
[16] Hurt E M, Kawasaki B T, Klarmann G J, Thomas S B and Farrar W L 2008 CD44 + CD24(-) prostate cells are early cancer progenitor/stem cells that provide a model for patients with poor prognosis *Br. J. Cancer* **98** 756–65
[17] Zola H *et al* 2005 CD molecules 2005: human cell differentiation molecules *Blood* **106** 3123–6
[18] Luo W, Tapolsky M, Earley K, Wood C G, Wilson D R, Logothetis C J and Lin S H 1999 Tumor-suppressive activity of CD66a in prostate cancer *Cancer Gene Ther.* **6** 313–21
[19] Beucher S 1994 Watershed, hierarchical segmentation and waterfall algorithm *Mathematical Morphology and Its*

Applications to Image Processing (Computational Imaging and Vision vol 2) ed J Serra and P Soille (Dordrecht: Springer)
[20] Jones T R, Carpenter A and Golland P 2005 Voronoi-based segmentation of cells on image manifolds *Lect. Notes Comput. Sci.* **3765** 535–43
[21] Tricot S *et al* 2015 Evaluating the efficiency of isotope transmission for improved panel design and a comparison of the detection sensitivities of mass cytometer instruments *Cytometry* A **87** 357–68
[22] Finck R, Simonds E F, Jager A, Krishnaswamy S, Sachs K, Fantl W, Pe'er D, Nolan G P and Bendall S C 2013 Normalization of mass cytometry data with bead standards *Cytometry* A **83** 483–94
[23] Lazar D C, Cho E H, Luttgen M S, Metzner T J, Uson M L, Torrey M, Gross M E and Kuhn P 2012 Cytometric comparisons between circulating tumor cells from prostate cancer patients and the prostate-tumor-derived LNCaP cell line *Phys. Biol.* **9** 016002
[24] Heinlein C A and Chang C 2004 Androgen receptor in prostate cancer *Endocr. Rev.* **25** 276–308
[25] Brooimans R A, Kraan J, van Putten W, Cornelissen J J, Lowenberg B and Gratama J W 2009 Flow cytometric differential of leukocyte populations in normal bone marrow: influence of peripheral blood contamination *Cytometry* B **76** 18–26
[26] Mohme M, Riethdorf S and Pantel K 2017 Circulating and disseminated tumour cells—mechanisms of immune surveillance and escape *Nat. Rev. Clin. Oncol.* **14** 155–67
[27] Gruber I, Landenberger N, Staebler A, Hahn M, Wallwiener D and Fehm T 2013 Relationship between circulating tumor cells and peripheral T-cells in patients with primary breast cancer *Anticancer Res.* **33** 2233–8
[28] Wistuba-Hamprecht K *et al* 2017 Establishing high dimensional immune signatures from peripheral blood via mass cytometry in a discovery cohort of stage IV melanoma patients *J. Immunol.* **198** 927–36
[29] Howat W J, Lewis A, Jones P, Kampf C, Ponten F, van der Loos C M, Gray N, Womack C and Warford A 2014 Antibody validation of immunohistochemistry for biomarker discovery: recommendations of a consortium of academic and pharmaceutical based histopathology researchers *Methods* **70** 34–8
[30] Uhlen M *et al* 2016 A proposal for validation of antibodies *Nat. Methods* **13** 823–7

III. Book chapter: Methods in Molecular Biology, Tumor Profiling

Thiele J-A, Pitule P, Hicks J and Kuhn P., **Single-Cell Analysis of Circulating Tumor Cells**. Methods Mol Biol., accepted in March 2018.

(Scopus cite score$_{2017}$=0.96)

Methods in Molecular Biology

Tumor Profiling

Chapter: **Single-Cell Analysis of Circulating Tumor Cells**

Jana- A. Thiele[1], Pavel Pitule[1], James Hicks[2,3] and Peter Kuhn[2,4]

[1]Biomedical Center, Faculty of Medicine in Pilsen, Charles University, Alej Svobody 76, 32300 Pilsen, Czech Republic.

[2]The Bridge Institute, Dornsife College of Letters, Arts and Sciences, University of Southern California, 3430 S. Vermont Ave., TRF 114, Los Angeles, CA 90089, USA.

[3]Cold Spring Harbor Laboratory, 1 Bungtown Rd., Cold Spring Harbor, NY 11724, USA.

[4] Department of Biomedical Engineering, Viterbi School of Engineering, University of Southern California, Los Angeles, California 90089

Corresponding author

Peter Kuhn: peter.kuhn@usc.edu
The Bridge Institute, Dornsife College of Letters, Arts and Sciences, University of Southern California, Los Angeles, CA 90089, USA.

Running Head: Single-Cell Analysis of CTCs

Abstract

Circulating tumor cells (CTCs) are rare cells, that can be found in the peripheral blood of cancer patients. They have been demonstrated to be useful prognostic markers in many cancer types. Within the last decade various methods have been developed to detect rare cells within a liquid biopsy from a cancer patient. These methods have revealed the phenotypic diversity of CTCs and how they can represent the complement of cells that are found in a tumor. Single-cell proteogenomics has emerged as an all-encompassing next-generation technological approach for CTC research. This allows for the deconstruction of cellular heterogeneity, dynamics of metastatic initiation and progression, and response or resistance to therapeutics in the clinical settings. We take advantage of this opportunity to investigate CTC heterogeneity and understand their full potential in precision medicine.

The High-Definition Single-Cell Analysis (HD-SCA) workflow combines detection of the entire population of CTCs and rare cancer related cells with single-cell genomic analysis and may therefore provide insight into their sub-populations based on molecular as well as morphological data.

In this chapter we will describe in detail the protocols from isolation of a candidate cell from a microscopy slide, through whole genome amplification and library preparation, to CNV analysis of identified cells from the HD-SCA workflow. This process may also be applicable to any platform starting with a standard microscopy slide or isolated cell of interest.

Keywords

Circulating tumor cells, CTC, copy number variation, CNV, liquid biopsy, precision medicine, single-cell analysis

1. Introduction

Throughout the initiation and progression of carcinomas, cells are actively or passively released from a tumor into the blood stream. Often tumor associated cells detach during the process of vascularization, or tumor derived cells intravasate through destabilized cell junctions into a blood vessel ***(1–4)***. These possible messengers of metastasis then circulate in the blood of patients and have been correlated with disease progression and patient outcome in numerous studies ***(5–7)***. The variety of tumor and tumor associated cells and extracellular vesicles (EVs) that pass over into the blood stream include circulating tumor cells (CTCs), endothelial cells which line the tumor vasculature, stromal cells, platelets, exosomes, oncosomes (oncogene-containing EVs) and more ***(8–10)***.

The detection of CTCs can be achieved by various methods ***(11, 12)*** utilizing mostly specific biological or physical properties of known cell-types, to enrichi a sub-set of CTCs ***(13–15)***. Such selection process can perform exceedingly well if the selection parameters are known with high accuracy and can be chosen with high precision. This process can be useful in a known diagnostic framework but can be challenging in a heterogenous disease setting represented by carcinomas. A possible solution is the direct analysis of the entire population of peripheral blood mononuclear cells (PBMC) from a patient to allow the unbiased identification of all CTC types. Deeper molecular analysis of the identified candidate cells can follow without losing valuable information especially of rare sub-populations that might be responsible for relapse or disease progression.

One technical solution to the challenge of hetereogeneity is the High-Definition Single-Cell Analysis (HD-SCA) workflow that has been developed to take the entire population of nucleated circulating cells in a peripheral blood sample into account ***(16, 17)***. It offers an enrichment-free, high-throughput approach for rare cell detection and multimodal single cell analysis following the principle that ‘no cell is left behind’ to be able to identify the one needle that matters in the haystack. For deeper insight about the impact of the detected rare cell, it is important to consider the spatiotemporal evolution of cancer; the migration of a CTC through the hematogeneous and lymphatic system of the body and

changes it might provoke within a new metastatic site ***(18)***. Collecting various available spatial and temporal data provides the opportunity to detect patterns and may predict patient outcome.

Recently the HD-SCA workflow has been extended to a variety of tissue preparations such as 'touch-preps'***(19)*** where resected tumor tissue or biopsy specimens are lightly touched to the surface of a slide, leaving an imprint of the tissue comprising a layer of intact tumor and associated cells that maintains their positional information. Being able to not only investigate peripheral blood and bone marrow aspirate, but also solid tumor tissue cells from the primary tumor and metastases, allows the HD-SCA workflow to correlate high-resolution imaging and downstream molecular data of single-cells of different spatiotemporal samples of the same patient to fully monitor and analyze disease progression.

This chapter aims to describe the methods for extracting individual cells of interest (COIs) from microscopy slides and to prepare them for genomic analysis, resulting in genome-wide copy number variation (CNV) profiles. This workflow is equally applicable not only to the PBMC fraction of a peripheral blood draw, but also to any slide-based cell preparation including 'touch-preps'. The detailed methods for the handling of blood samples, preparation for immunofluorescent staining and characterization of COIs from blood samples have been previously published ***(16)***, but will be briefly described here for the better understanding of the analytical steps of the workflow.

2. Materials

All materials are listed in order of time of use for the corresponding part of the protocol listed below.

2.1 High-Definition Single-Cell Analysis (HD-SCA) workflow

1. Blood collection tubes (BCTs) containing a cell fixation and DNA stabilization reagent (Streck Cell-Free DNA BCT®)
2. Customized adhesive glass slides (Marienfeld GmbH & Co. KG)
3. Fluorescently stained antibodies (AlexaFluor©, Thermo Fisher Scientific)

2.2 Single-cell isolation

2.2.1 Equipment

1. Inverted microscope, e.g. Olympus IX81, Nikon TE2000
2. Micromanipulator, e.g. TransferMan® 4r (Eppendorf)

2.2.2 Reagents and consumables

3. DNA AWAY® surface decontaminant (Molecular BioProducts)
4. Cell isolation cover slips: 25x75 mm (Electron Microscopy Sciences)
5. Glass capillaries (piezo drill tips): diameter of 15µm, angle of 25°, lengths of 6.000 µm and a jagged front surface (Eppendorf)
6. Oil for hydraulic system: mineral oil
7. Dry ice and wet ice to store cells and buffers (in thermal boxes)
8. Picking buffer PBS-T: 5 ml 1X PBS (pH=7.4 ±0.02) + 5 µl of Tween 20 (0.1 % final concentration of Tween 20)
9. Cell deposition buffer (CDB): 10 mM Tris-HCl-EDTA pH 8.0 (TE): 10 mM Tris-HCl, 1 mM disodium EDTA, pH 8.0
10. 0.5 ml LoBind DNA tubes: sterile, with flat rim (Eppendorf)
11. 70 % Ethanol
12. Compressed air

2.2.3 Software and computing

13. Imaging Software: (ImagePro, MediaCybernetics Inc.)
14. Semi-automated custom ImagePro macros for relocation and imaging of cells
15. Transformation matrix: Converts slide scanner coordinates for each cell of interest to reimaging/cell picking microscope coordinates.

2.3 Whole genome amplification

2.3.1 Equipment

1. Thermal cycler (e.g. Mastercycler™ pro PCR System, Eppendorf)
2. PCR cooler (e.g. iceless cold storage system, Eppendorf®)
3. Equipment and chemicals for gel electrophoresis system: e.g. Quick-Load® 100bp DNA Ladder (New England BioLabs), 10X SYBR® Safe DNA Gel Stain (Life Technologies), Gel Pilot Loading Dye 5X (Qiagen), GelPilot® Agarose (Qiagen)
4. DNA quantification instrument (e.g. Qubit, ThermoFisher Scientific)

2.3.2 Reagents and consumables

5. DNA AWAY® surface decontaminant (Molecular BioProducts)
6. 70 % EtOH
7. Lysis buffer: 1:1 dithiothreitol (DTT, 100 mM) + potassium hydroxide (KOH, 400 mM)
8. GenomePlex® Single Cell Whole Genome Amplification Kit (WGA4, Sigma-Aldrich): 10X single cell lysis & fragmentation buffer, 1X single cell library preparation buffer, library stabilization solution, library preparation enzyme, 10X Amplification Master Mix, WGA DNA polymerase, Control gDNA (5 ng/ml), molecular grade water
9. 10 mM Tris-HCl-EDTA pH 8.0 (TE): 10 mM Tris-HCl, 1 mM disodium EDTA, pH 8.0
10. QIAquick PCR Purification kit (Qiagen)
11. Qubit dsDNA HS Assay Kit (ThermoFisher Scientific)
12. Qubit assay tubes (ThermoFisher Scientific)

2.4 Library Construction

2.4.1 Equipment

1. Sonication device, e.g. Covaris S2 (COVARIS Inc.)
2. Thermal cycler (e.g. Mastercycler™ pro PCR System, Eppendorf).
3. DNA quantification instrument (e.g. Qubit, ThermoFisher Scientific)
4. Bioanalyzer; e.g. 2100 Bioanalyzer (Agilent Technologies)
5. Magnetic stands for PCR tubes/96well plates (e.g. LifeTechnologies)

6. Magnetic stand for 1.5 ml tubes (e.g. DynaMag®, ThermoFisher Scientific)

2.4.2 Reagents and consumables

7. Sonication tube for fragment size of 200-250 bp, e.g. Snap Cap microTUBE (COVARIS Inc.)
8. 10 mM TE (pH= 7.5-8.0, RT): 10 mM Tris-HCl, 1 mM disodium EDTA
9. DNase free 0.2 ml PCR 8-strips
10. NEBNext® Ultra DNA Library Prep Kit for Illumina® (New England BioLabs Inc.)
11. NEBNext® Multiplex Oligos for Illumina®-Dual Index Primer Set 1 (New England BioLabs Inc.)
12. Agencourt® AMPure® XP Beads (Beckman Coulter)
13. 100 % EtOH
14. DNA low-bind tubes (Eppendorf): PCR tubes, 1.5 ml tubes, 96 well plates
15. Qubit dsDNA HS Assay Kit (ThermoFisher Scientific)
16. Qubit assay tubes (ThermoFisher Scientific)
17. High Senitivity DNA Analysis Kit (Agilent Technologies)

2.5 *Single-Cell CNV profiling analysis software tools*

1. Bowtie (sequence analysis software; 2.2.6 or later)
2. R: The R project for statistical computing (or equivalent)
3. University of California Santa Cruz (UCSC) Genome browser: https://genome.ucsc.edu/, hg19 reference genome

3. Methods

The entire procedure from initial blood sample preparation to sequence-ready DNA library takes approximately 5 days of elapsed time, is visually summarized in **Figure 1** and consists of the following major steps (the steps described in detail in this chapter are marked in bold letters):

1. *Blood sample collection (time to sample process of up to 48 hours to enable standard shipping conditions)*
2. *Blood sample processing (half day of elapsed time, 45 minutes of hands-on/sample)*
3. *Fluorescent staining (half day of elapsed time, 5 minutes of hands-on/slide)*
4. *Whole slide imaging (1.5 hrs of elapsed time/2 slides)*
5. *Technical analysis of COIs (5 min/slide)*
6. ***Single-cell isolation (3 min hands-on/cell)***
7. ***Whole genome amplification including cleanup (half day for 24 cells of elapsed time, 2h of hands-on/24 cells)***
8. ***Library construction including sonication (2 days for 24 cells of elapsed time, 5h hands-on /24 cells)***
9. ***Single-Cell CNV profiling***

Details for required steps to detect COIs on glass slides (steps 1-5 above) are provided in Marrinucci et al. ***(16)*** and are discussed briefly in section 3.1 below. The detailed protocol provided in sections 3.2-3.5 below describes the necessary steps from isolation of a single candidate cell to genome-wide CNV analysis.

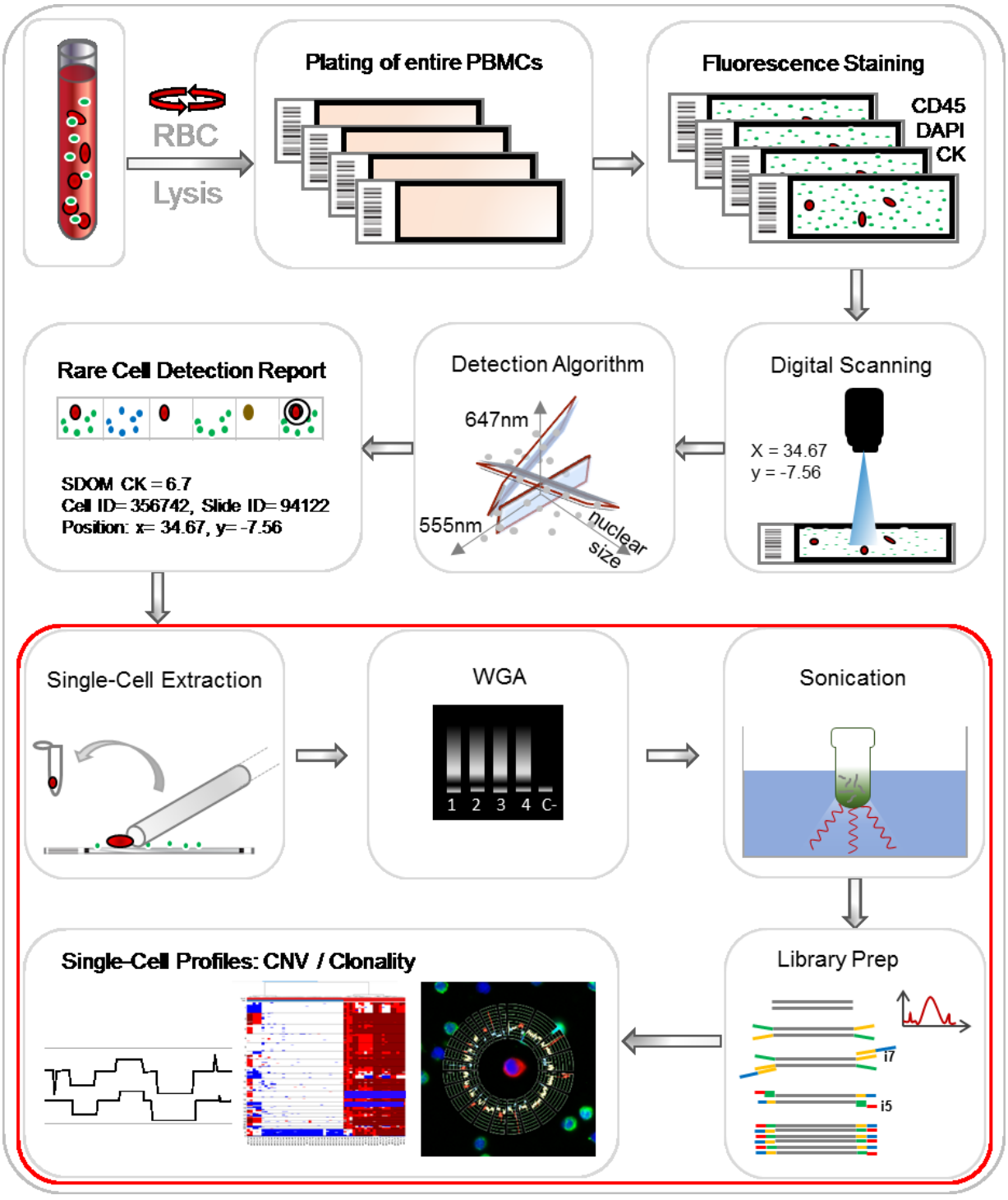

Figure 1. Overview of the complete HD-SCA workflow. The HD-SCA workflow from a blood draw to a single-cell CNV profile: Red blood cell (RBC) lysis is followed by plating of the entire PBMC fraction, fluorescent staining, full slide scanning and detection of COIs. The CTC detection is followed by single-cell extraction, next-generation sequencing and CNV profiling. **In red:** steps provided within this protocol: Single-cell extraction through micromanipulation, WGA, library preparation (including sonication) and single-cell CNV profile analysis.

3.1 Summary of the HD-SCA workflow

Blood is drawn in blood collection tubes (BCTs) containing a cell fixation and DNA stabilization reagent and then shipped in temperature controlled boxes to ensure preservation of intact cells and DNA. Samples are processed by plating of the PBMC fraction on customized adhesive glass slides after red blood cell (RBC) lysis.

Slides are then stained for immunofluorescent identification of rare cells of interest among the white blood cell (WBC) population. Briefly, cells are fixed, permeabilized and afterwards fluorescently stained with antibodies against a pan-cytokeratin panel (targeting an epithelial-specific intermediate filament), CD45 (a leukocyte specific marker) and DAPI (4',6-diaminido-2-phenylindole, a nucleic acid stain). The markers for identification can be adapted depending on the COI, hence specialized assays e.g. for melanoma cells ***(20)*** and endothelial cells ***(21)*** are available. An additional disease-specific marker can be used as a forth channel marker to study the biology of the CTC. Fixation steps (chemically in the tube, physically on the slide and during staining) are not only necessary in order to identify potential CTCs, but also present a challenge, especially for the quality of downstream analysis ***(22)***. The immunofluorescence staining is followed by a high-throughput digital imaging pipeline. All images are stored and analyzed by an R-based software routine.

Taken into consideration by the analysis algorithms are morphology data like nuclear size or shape and signal intensities or absence of the epithelial and leukocyte marker to identify cells distinct from a WBC. Finally the cells that have been calculated as rare events and probable high definition circulating tumor cells (HD-CTCs; $CK^{pos}/CD45^{neg}$ with distinct nuclear shape) are presented to a hematopathologically trained specialist for final technical analysis and classification. An example for different categories of CTCs analyzed by the HD-SCA workflow are displayed in **Figure 2**. The final result is an enumeration of all rare events within the entire PBMC cell population and a comprehensive morphometric data set for each cell.

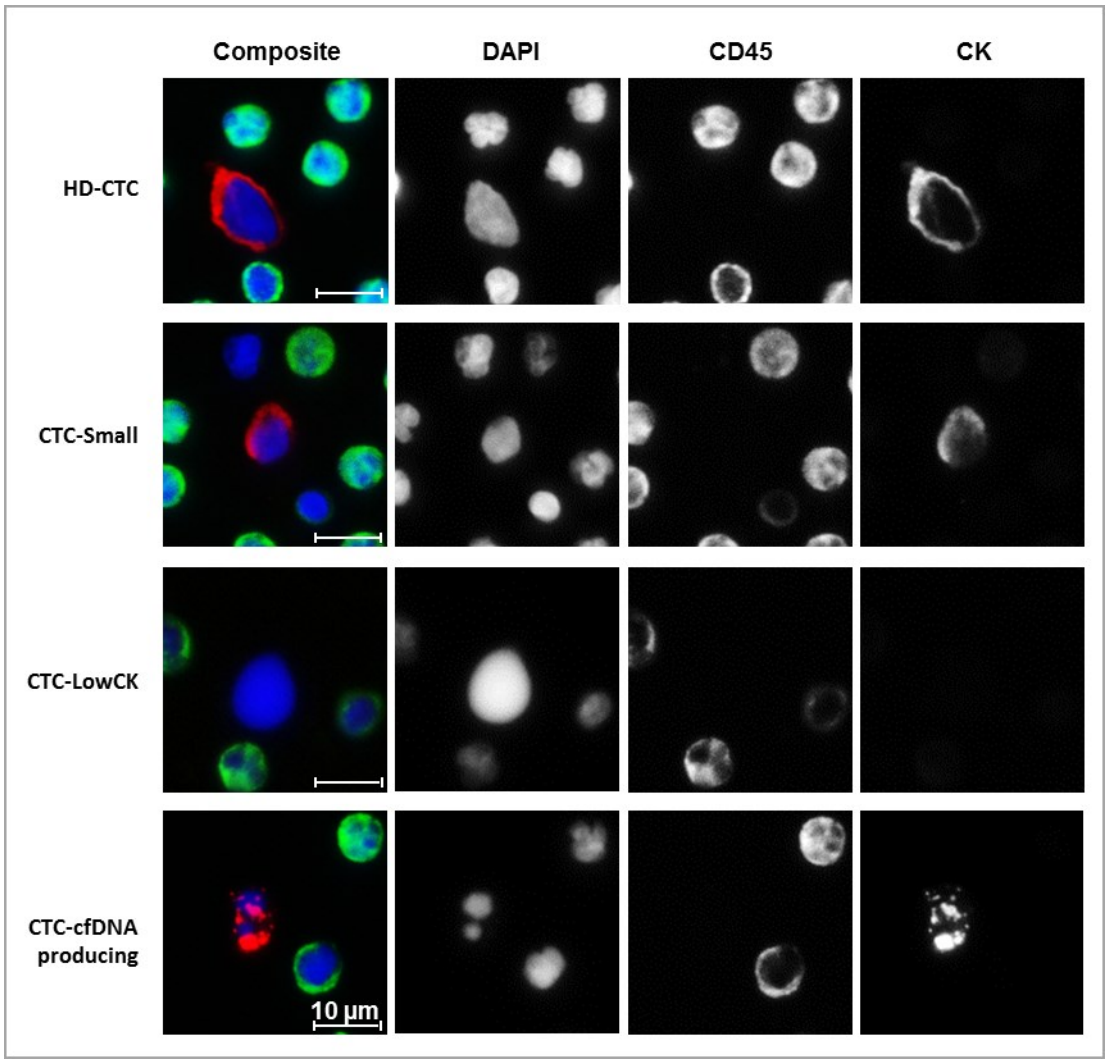

Figure 2. Categories of detected potential CTCs. These CTCs are representing the pleomorphic character of CTCs detected in the blood of a metastatic prostate cancer patient. Cells have been stained for nuclei (DAPI; blue), CD45 (green) and CK (red). Displayed are the composite and the single channel images. **HD-CTC:** $CK^{pos}/CD45^{neg}$ cells with a nuclear shape distinct from WBCs. **CTC-Small:** $CK^{pos}/CD45^{neg}$ cells with a small nucleus (WBC-like). **CTC-LowCK:** $CK^{neg}/CD45^{neg}$ cells with a nucleus at least double the size of a WBC. **CTC-cfDNA producing:** CTCs undergoing apoptosis ($CK^{pos}/CD45^{neg}$ with irregular cytoplasmic or nuclear condensation) and about to release circulating tumor DNA. Scale bars equal 10 µm.

For all cells on a slide, approximately 400 variables are detected and stored, including the position of each COI on the slide by their Cartesian coordinates. Two numerical coordinates define the exact position on the slide, using two perpendicular lines as reference axes. This coordinate system can be applied to any microscope without special instrumentation and allows the relocation of each cell further along the pipeline. In order to perform single-cell downstream analysis, COIs are relocated and imaged at high definition to enable subsequent correlation of genomic and morphological data. COIs are typically imaged at 400X magnification on a fluorescence microscope, but may also be analyzed

through deconvolution or confocal microscopy. Each COI can then be isolated from the slide using the Cartesian coordinates and a micromanipulation station with a glass capillary and a hydraulic cylinder system.

3.2 Single-Cell Isolation

After a COI has been identified, cell isolation is started, but before isolation from the slide, reimaging of COIs at 400x magnification is recommended to achieve a high quality image. This enables future cytomorphologic downstream analysis of COIs.

1. Prepare all buffers and keep them on ice and check oil level in micromanipulation system (if you intend to proceed with cell isolation after relocation and reimaging of the cell).
2. Clean slide with 70 % ethanol (do not disturb coverslip) and remove dust with compressed air spray.
3. COI is relocated using the imaging software ImagePro and a custom macro that uses a transformation matrix, which maps the coordinates between the scanning and reimaging microscope. 100X images from the scanning report are used to confirm the location. Detected offset can be applied to all other coordinates of COIs on this slide.
4. Each cell position is confirmed and images are taken in each fluorescent channel using a 400X magnification until all COIs of the slide are imaged.
5. Color composites and images of individual channels are stored in a database.

Pause point: Slides can be stored in a dark and dry location for future downstream analysis.

6. To proceed with cell isolation, clean all areas around the microscope station with DNA AWAY® and 70 % ethanol. Place a cover slip on the slide holder insert to provide PCR tube support. Prepare 0.5 ml PCR tubes using UV sterilization prior to use.
7. Peel off nail polish from slide carefully without moving the cover slip (a scalpel may be useful).
8. Place slide in 1X PBS in coplin jar in a tilted position until cover slip comes off (~10 min).
9. Place slide (without coverslip) on microscope next to empty cover slip and add 1 ml of PBS-T to prevent cells from dehydration (refill if PBS-T starts to evaporate during longer cell extraction durations).

10. Insert glass capillary into the micromanipulation arm. Ensure straight orientation of capillary.
11. Use microscope software control box to turn on the live preview and choose brightfield with a 20X objective lens; use stored coordinates to navigate to cell.
12. Use control panels (one for oil pressure, one for capillary, one for microscopy stage) to navigate end of capillary in field of vision, but keep capillary above slide surface, but within PBS-T.
13. Focus on capillary tip and adjust oil so that air (visible as dark mass) in the capillary moves close to tip opening.
14. Focus on cell and slowly navigate capillary down and in front of cell as described and illustrated in **Figure 3**.
15. Dislodge COI using capillary tip or use the tip to push other cells out of the way first (*see* **Note 1**).
16. Position capillary tip right in front of loose COI and use oil pressure wheel to create suction and aspirate the COI into the capillary (make sure to ONLY aspirate the COI, not other cells!). Move capillary up (out of the PBS-T); do NOT move capillary in x and y direction!
17. Place a PCR tube with open lid (opaque label facing up) on the cover slip next to the slide and place a 1 µl drop of cell deposition buffer (CDB) between the first two lines behind the opening of the tube.
18. Now move microscope stage to steer the tube to the capillary tip to position CDB below tip. Do not touch the capillary tip with the tube.
19. Focus on buffer drop (edge of drop has to be a sharp, dark outline) and move tip down until it is immersed in the drop.
20. Focus on tip and use oil wheel to slowly release cell. Observe tip opening, watch the cell moving through the capillary and make sure it is in the drop.

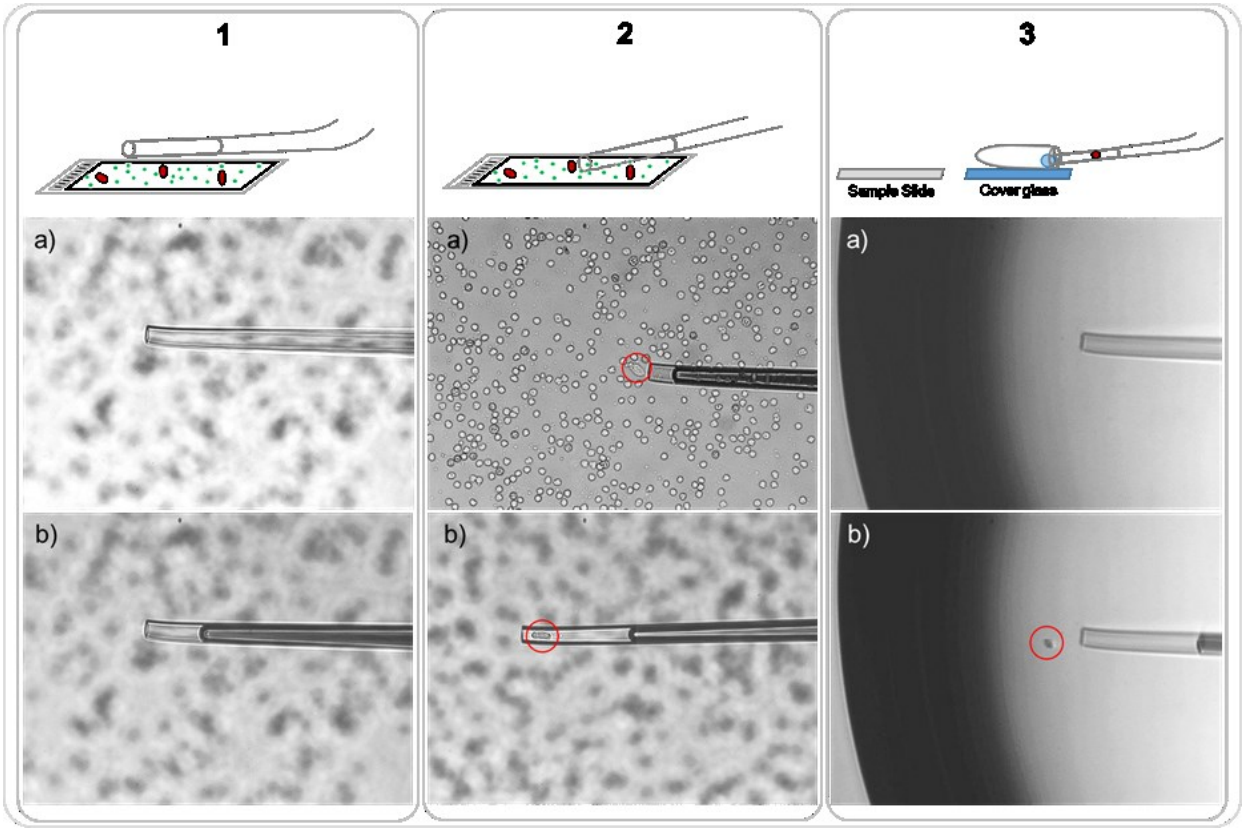

Figure 3. The Single-Cell Isolation Process. 1) The capillary is positioned inside the buffer solution above the cell of interest and **1a)** set into focus. **1b)** The oil adjustment wheel is used to move the air close to the capillary opening and then stabilized there. **2)** The cell of interest (COI) is set into focus and the capillary tip moved down onto the slide **2a)** to be in focus together with the COI. The capillary is used to loosen the COI and the oil wheel to create suction to **2b)** aspirate the cell into the capillary. The capillary is moved up, out of the buffer. **3)** A PCR tube is placed on a cover slip next to the sample slide. **3a)** The edge of the 1 µl drop of TE buffer is set into focus and the capillary is lowered into the drop. **3b)** The COI is released carefully by slowly turning the oil wheel to the "out" direction.

21. Move tip up and the stage away from tip. Take the tube, and tap slightly on bench to let the drop with the cell move towards the bottom before closing the lid. Follow with a quick spin and store tube directly on dry ice.
22. Mark cell in your software or database as 'isolated' and move to next cell position; repeat steps 11-20.

3.3 Whole genome amplification

Following the extraction of single cells, whole genome amplification (WGA) and fragmentation by sonication is used to reach a sufficient amount of DNA in the required fragment length for single-cell sequencing. Since a diploid human cell only contains about 7 pg of genomic DNA, amplification is

necessary prior to genetic variation analyses ***(23)***. An overview of all procedures involved in WGA and library preparation with time estimates are indicated in **Figure 4**.

1. Clean all surfaces with DNA AWAY® and 70% EtOH.
2. Thaw PCR tubes containing the single cell in 1 µl of TE, always add a sample for negative control (TE only) and positive control (1.5 µl gDNA).
3. Add 1.5 µl of lysis buffer (1:1 DTT & KOH) & spin down (gently).
4. PCR program: 95°C for 2 min.
5. Cool tubes on a PCR cooler and prepare master mix:
    - Add 6.5 µl of 10 mM TE for each reaction
    - Add 1 µl 10X Single Cell Lysis & Fragmentation Buffer per reaction, mix thoroughly and spin down

    Add 7.5 µl of master mix to each reaction, including controls.
6. PCR program: 99°C for 4 min (time sensitive: take samples out instantly and place on PCR cooler).

Pause point: Store the reactions at -20°C or keep on a PCR cooler to continue immediately.

7. Prepare master mix of:
    - 2 µl of 1X Single Cell Library Preparation Buffer
    - 1 µl of Library Stabilization Solution per reaction and mix thoroughly

    Add 3 µl of the mix to each reaction.
8. Mix thoroughly and place in thermal cycler at 95°C for 2 min, cool samples on PCR cooler, spin down and store on PCR cooler.
9. Add 1 µl of Library Preparation Enzyme to each reaction, mix thoroughly and spin down.
10. Place samples in thermal cycler and incubate as follows:

    16°C for 20 min, 24°C for 20 min, 37°C for 20 min, 75°C for 5 min, 4°C hold.

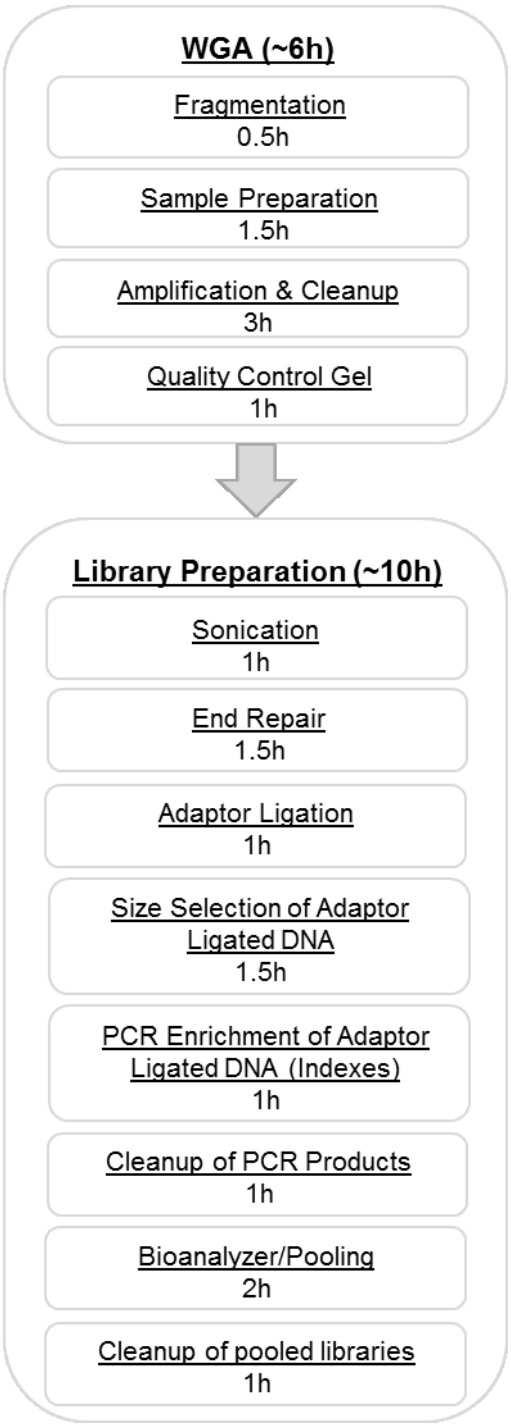

Figure 4. Whole genome amplification (WGA) and library preparation workflow chart. Overview and approximate timing estimation of all steps included in the process from extracted single cell to ready-to-sequence library. Time estimates are based on processing of 24 samples.

11. Spin samples down.

Pause point: Store at -20°C for up to three days or amplify immediately.

12. Prepare master mix of:

- 7.5 µl of 10X Amplification master mix
- 48.5 µl of molecular grade water (provided in the kit)
- 5 µl of WGA DNA Polymerase

per reaction and mix thoroughly. Add 61 µl of the mix to each reaction.

13. Mix thoroughly, spin down and place in thermal cycler:

    95°C for 3 min, 24 cycles of: 94°C for 30 sec, 65°C for 5 min; then hold at 4°C.

Pause point: Store the reactions at -20°C (pause point) or keep at 4°C to continue immediately.

14. Prepare a 1.5 % agarose gel including gel stain. Mix 2 µl 5X loading dye with 8 µl DNA sample and load samples onto agarose gel. Add Quick-Load® 100 bp DNA Ladder in one well and run for 35 min at 90 V.
15. Take image of agarose gel and evaluate samples: Successful WGA produces a continuous smear mainly between 150 and 1000 bp. If no smear is visible, WGA was not successful and sample cannot be used for library preparation (see **Note 2** and **Figure 7**)
16. Use QIAquick PCR Purification kit according to manufacturer instructions.
17. Keep eluted DNA on PCR cooler and quantify all samples using e.g Qubit quantification system according to manufacturer instructions; note DNA concentrations in an excel sheet or similar form.

3.4 Library Construction

Before starting library construction all samples have to be adjusted to same concentrations to obtain equal number of reads per sample during sequencing.

1. For each sample use 185 ng DNA input in 55.5 µl total volume, therefore calculate your sample volume and add 10mM TE buffer accordingly.
2. Transfer all samples to a sonication tube suitable for your equipment, volume and fragment size (here: a COVARIS Snap Cap microTUBE).
3. Follow the manufacturers protocol to reach a fragment size of approximately 200 – 250 bp and continue library preparation with fragmented DNA.
4. Transfer the fragmented DNA samples to strips of 0.2 ml PCR tubes.

5. Prepare a master mix using the NEBNext End Prep Kit:
    - 3 µl End Prep Enzyme Mix
    - 6.5 µl 10X End Repair Reaction Buffer

    Add 9.5 µl of mastermix to each 55.5 µl fragmented DNA, mix by pipetting and spin down.
6. Place in a thermal cycler and incubate:

    20°C for 30 min, 65°C for 30 min, 4°C hold
7. Prepare master mix:
    - 15 µl Blunt/TA Ligase Master Mix
    - 2.5 µl NEBNext Adaptor for Illumina
    - 1 µl Ligation enhancer

    Add the 18.5 µl master mix immediately to the 65 µl of cooled down reaction mix. Total volume is 83.5 µl.

    Mix by pipetting and spin down.
8. Place in a thermal cycler and incubate at 20°C for 15 min, then place on PCR cooler.
9. Add 3 µl of USER enzyme to each sample reaction mix.
10. Spin down and incubate at 37°C for 15 min.

Pause point: Store the reactions at -20°C or keep on PCR cooler to continue.

11. Thaw reagents for step 12 for 30 min, keep DNA on PCR coolers until use.
12. For AMPure XP bead size selection of adaptor ligated DNA, follow the manufacturer instructions to select fragments between 200 and 300 bp. Let beads and TE come to RT, vortex beads, prepare fresh 80% ethanol.

    At the end of the AMPure XP bead size selection protocol, elute in 17 µl TE and finally transfer 15 µl to a new PCR tube for amplification.
13. For PCR Enrichment of Adaptor Ligated DNA mix the following components in fresh sterile nuclease-free tubes (no master mix here!). Make sure to record index numbers for each sample and to assign to each reaction a unique combination of indexes within one set of libraries that is about to get pooled:

- 15 µl Adaptor Ligated DNA Fragments
- 25 µl NEBNext Q5 Hot Start HiFi PCR Master Mix
- 5 µl Index Primer/i7 Primer
- 5 µl Universal PCR Primer/i5 Primer
- Total reaction volume: 50 µl

14. Place in thermal cycler and run the following program:

    98°C for 30 sec; 7 cycles of: 98°C for 10 sec, 65°C for 75 sec; then final extension: 65°C for 5 min and hold at 4°C.

15. Use AMPure XP beads for PCR product cleanup: let beads and TE come to RT, vortex beads, prepare fresh 80 % ethanol.
16. Transfer library into low-binding 1.5 ml tube, add 0.8 volume of beads (for 50 µl sample: 40 µl beads) to sample, mix thoroughly, but gently by pipetting up and down and incubate 5 min at RT.
17. Spin down, place in magnet stand with open lid, incubate 5 min at RT, discard supernatant.
18. Wash twice with 200 µl of 80% EtOH, incubate 30 sec, discard supernatant.
19. Air dry beads for exactly 2 min with lid open on magnet.
20. Remove tube from magnet and elute DNA in 35 µl TE, mix well, incubate for 3 min at RT, spin down, place in magnet, let sit for 5 min.
21. Collect 32 µl supernatant, that contains the DNA, discard beads.
22. Measure DNA quantity (e.g. Qubit); following the manufacturer's protocol.
23. Use a Bioanalyzer for analysis of size distribution by following the manufacturer's instructions.
24. Quality control of Bioanalyzer size distribution check: If the library preparation was successful, the lower and upper marker are framing a normal curve of distribution with a peak around 300-400 bps (see **Note 3**). A low concentration of primer-dimers and adaptor dimers (<130 bps) might be observed right after the lower marker, but these low concentrations will be eliminated during the final cleanup.

    If single peaks are observed, especially in the area of short fragments (primer-dimer: ~80 bps or adaptor-dimer: ~130 bps) and no fragment distribution with a peak around 380 bps, library preparation was not successful. Examples for libray control results are displayed in **Figure 5**.

25. Calculate the molarity of each library:

$$library\ nM\ = \frac{\left[Qubit\ value\ (\frac{ng}{ul})\right] * 1000}{1} * \frac{1}{649\ \left(\frac{M}{bp}\right)} * \frac{1000}{size\ of\ library\ peak\ (bp)}$$

Pool all libraries at a final concentration of 10 mM using a volume of 5 µl/sample (dilute with TE buffer).

26. Perform one last cleanup using the AMPure XP beads (repeat steps 15-18) with the total volume of all pooled libraries: $Total\ Vol. = Nr. of\ libraries * 5\mu l$.
27. Air dry beads for 2 min with lid open, elute DNA in the same volume of TE as the volume you started the cleanup with ($Total\ Vol.$). Mix well, incubate for 3 min at RT, spin down, place in magnet, let sit for 5 min.

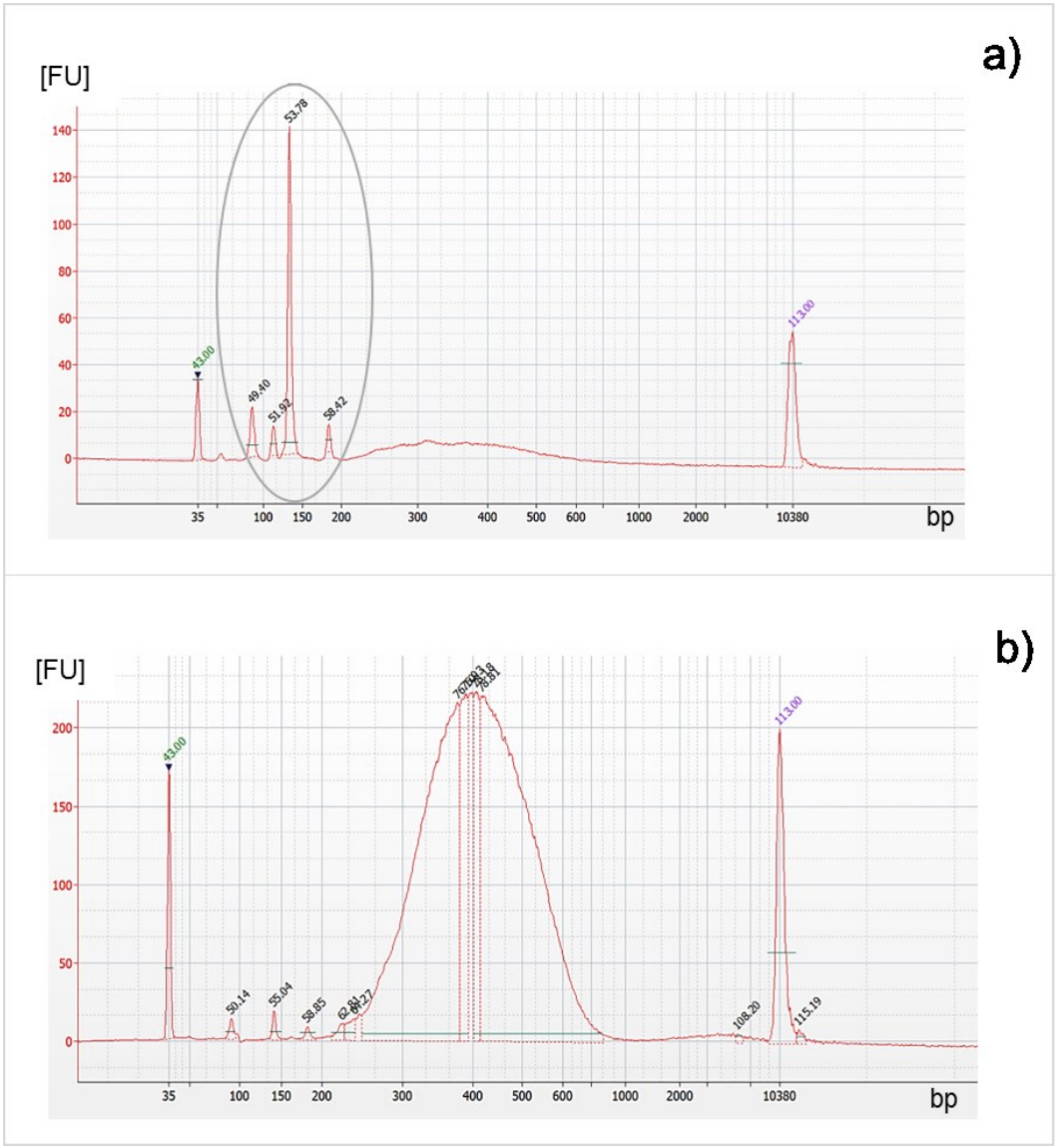

Figure 5. Quality control of size distribution after library preparation using the Bioanalyzer. a) An example of a poor quality library with primer dimers (~80 bp) and a high number of adaptor dimers (~130 bp), both within grey encircled area. The concentration of the library is very low (represented by a fluorescent signal <10 FU), but the size (~350 bp) is correct. Even after an additional cleanup this library would not be sufficient for sequencing. **b)** Adaptor dimers and primer dimers are of very low concentration. Library shows a perfect size distribution with a peak around 380 bp and a high concentration (>200 FU). After the final cleanup, this library will be sufficient for next-generation sequencing.

28. Collect as much supernatant (that contains the DNA) as possible without touching the beads, discard beads.
29. Pooled library is ready for sequencing. Protocols for next generation-sequencing of libraries for CNV analysis are in detail described in Baslan et al. (here, DNA libraries were sequenced on the HiSeq2500 platform (Illumina) using single end read 50 base pair protocol (SR50)) ***(24)***.

3.5 Single-Cell CNV profiling

Single-cell CNV profiling has become a valuable tool in CTC research to monitor the evolution of disease during the course of therapy ***(25–27)***. It enables detection of new lineages or subclones appearing in the circulation that reflect changes in the tumor load that would not be easily detected by bulk based applications ***(27)***. As described by Dago et al. ***(25)*** such rare subclones may play an important role in chemotherapy resistance and disease progression and their early detection could thus greatly improve therapeutic outcomes.

Procedures based on the protocols of Baslan et al. ***(24)*** are then performed to allow CNV profiling of single CTCs, and this reference provides method details for single-cell CNV profiling including all processing steps, software details and scripts. An overview of the single-cell CNV profiling process is provided here. After sequencing, results are transferred as a fastq-file, the open-source software 'Bowtie' can be used to align the sequences to the reference genome 'hg19', which can be acquired from the UCSC Genome Browser. Before the sequences can be mapped, a set with an arbitrary number of 'bins' is created across the genome with each bin containing the same number of mappable positions (intervals of sequence reads). A high number of bins will allow for higher resolution but is also more sensitive for noise, which will lower confidence in observed events. A good balance between resolution and noise has to be determined and will differ on a cell to cell basis due to differences in whole genome amplification and sequencing efficiency. This concept of optimized sparse sequencing of single cells for copy number analysis has been developed by Navin et al. ***(28)***. In 2012 Baslan et al. ***(24)*** improved the method for high-throughput, but low cost single-cell sequencing.

The optimized method of sparse sequencing allows for accurate, but cost-efficient acquisition of CNV profiles even from single-cells. It requires 20 times fewer reads compared to Navin et al. ***(28)*** and only about 75 million bases instead of over 100 billion as is standard for whole genome sequencing ***(29)***.

Once the binned data has been acquired, it can be simplified using mathematical segmentation ***(30)***. Segmentation quantifies the amplitude and location of copy number gains and losses across the genome, creating a genome-wide, numerical 'CNV profile' of each cell. CNV profiles, in form of numerical vectors, can be clustered to identify the lineage relationships among cells in a blood draw or to compare and contrast cells from different timepoints or different patients. The degree of similarity among groups of individual CNV profiles can reveal clonal and subclonal structures in the population that are typical of cancer. Among CTCs, the identification of complex clonal CNV profiles distinguishes the cancer cells from white blood cells or other cells in circulation that exhibit 'flat' profiles typical of normal diploid genomes with no significant copy number alterations (see **Figure 6**).

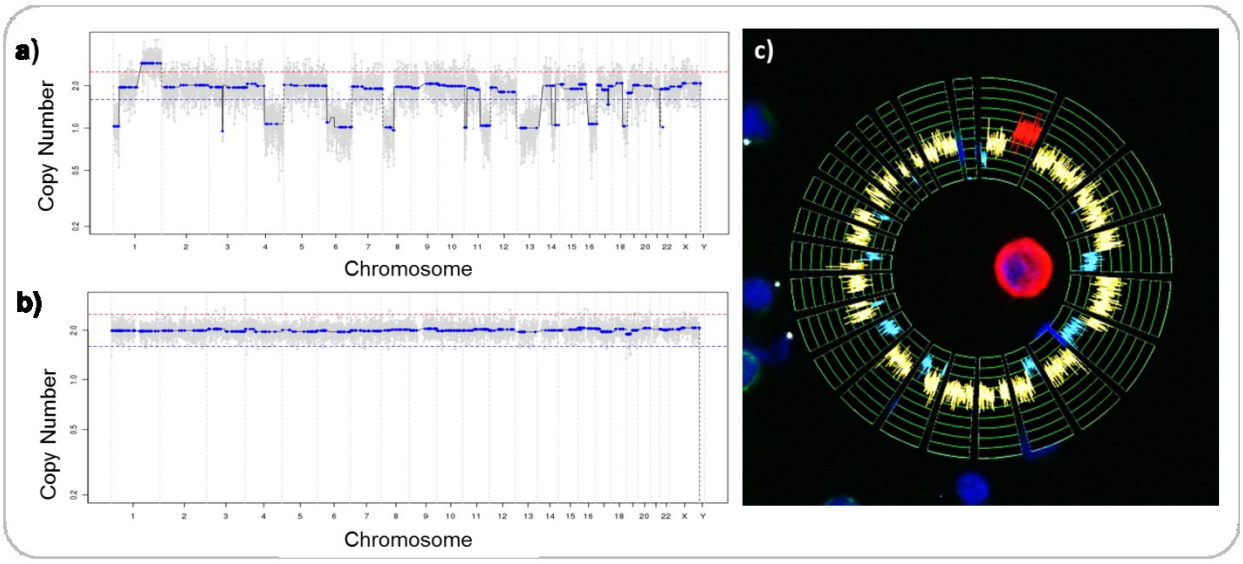

Figure 6. Copy Number Variation (CNV) profile examples. a) An example of a complex genome with multiple copy number alterations consistent with a malignant CNV profile of a COI. **b)** The profile of a diploid genome derived from a white blood cell (WBC) with no large copy number variation. **c)** Various graphic programs can be used to visually display CNV profiles. Here the software CIRCOS ***(31)*** has been used to illustrate CNVs in each chromosome of the cell displayed in the center, which is the same cell-profile as shown in a).

Most importantly, the tracking of single-cell CNV profiles across multiple timepoints during therapy provides an important 'window' into the genetic response of the cancer to treatment. Dago et al. ***(25)*** demonstrated that a cancer that was genetically stable for months can rapidly change its genomic make-up in response to treatment pressure. Although more work is clearly needed to understand how

to translate these genomic changes into improved outcomes, we believe that a combined approach of proteo-morphometric and genomic analysis on the rare cells identified in the HD-SCA workflow will be a key component of real-time cancer diagnosis and targeted therapies.

4 Notes

1 *Single-Cell Isolation troubleshooting*

- While using the capillary to push other cells out of the way or scraping the COI off, ensure to not have a cell stick to the capillary or aspirate loosened cells together with the COI.
- When aspirating the COI into the capillary, ensure to turn the oil wheel until the cell is all the way out of the field of vision (about two full turns), then one quarter-turn backwards to stop the oil flow.
- Not moving the capillary in x or y direction helps keeping the capillary in focus, enabling a faster process for depositing the COI in the CDB drop.
- If the cell 'disappeared' in the capillary during the attempt to place it in the drop, it can be caused by one of three issues:
    1. Low oil level in the system
    2. Air bubbles in the hose of the oil system
    3. Oil and air bubbles in the capillary

    Check oil level before isolating cells. Remove air bubbles from oil supply hose and make sure to have no air trapped in the hose while inserting the capillary. Change capillary after a maximum of about 10 cells to prevent oil moving into the capillary (danger of contaminating the sample).

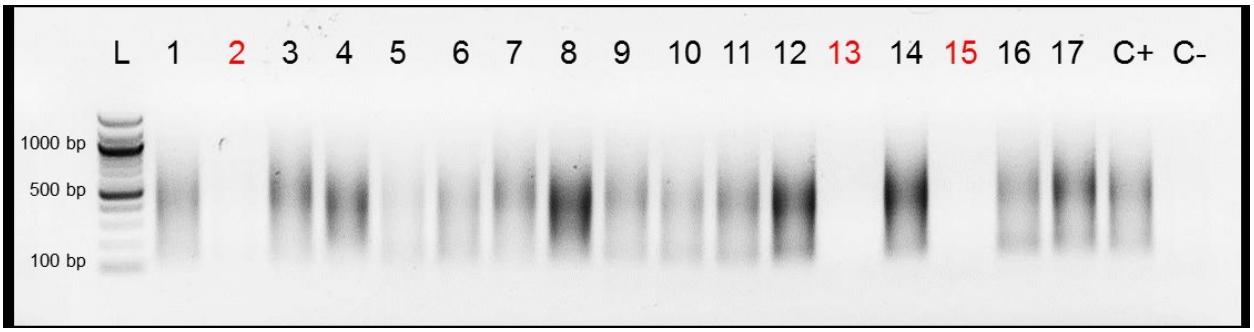

Figure 7. Whole Genome Amplification (WGA) control gel. On the left is the DNA ladder, followed by 17 samples, the gDNA control (C+) and the negative control (C-). The samples labeled in red (2, 13, 15) have no or only a very dim smear and therefore no sufficient amplification. In cells that show no smear, like in this example cells 2, 13, 15, cell transfer during isolation might have been incomplete or DNA quality of those cells insufficient.

2 *Whole genome amplification troubleshooting*

- If most samples show a smear, but a few do not (see **Figure 7**), it might be due to a loss of the cell during cell isolation. Make sure to observe the deposition of the cell into the buffer.
- If no smear is visible, but primer dimers are: PCR worked, but cells have not been lysed properly. Ensure that 10X Single Cell Lysis & Fragmentation Buffer it thawn and mixed well before use until no precipitation is visible.
- If neither smear nor primer dimers appear, the enzyme might not have worked sufficiently. Repeat 4 cycles of amplification PCR:

  95°C for 3 min, 4 cycles of: 94°C for 30 sec, 65°C for 5 min; hold at 4°C.

  Repeat agarose gel steps (3.2.: 14-15).

  If still neither smear nor primer dimers are visible, PCR has not worked and new reagents are required.
- If the negative control shows a smear, water and TE have to be discarded immediately. Other components of the kit should be tested (buffers and enzymes) again with a new negative and positive control to make sure the contamination is eliminated before moving forward with valuable cells. If contamination continues, all used components have to be discarded.
- If the positive control shows no smear, test the gDNA concentration using the Qubit. Concentration should be 5 pg/µl. Mix gDNA well before use.

3 *Library Construction troubleshooting*

- If amount of library is insufficient it could be caused by two things: Low quality or concentration of starting material. DNA input can be increased up to the maximum input in 55 µl sample volume (without additional water). Due to some volume loss during sonication, 60 µl can be used for sonication if the sonication protocol and tube size allows it. The quality

of the starting DNA (WGA products) can be determined via Bioanalyzer. In addition, make sure calculations were correct, and that magnetic beads and TE were brought to RT.

- To ensure that the final cleanup has successfully removed most primer and adaptor dimers, re-test all final libraries on a Bioanalyzer.
- If the Bionalayzer does not show a peak size of 300-400 bp (with a total library fragment range between 100 and 600 bp) the magnetic bead cleanup was not successful. During the cleanup ensure all components are brought to RT and that beads are vortexed well right before use.

References

1. Ashworth TR (1869) A case of cancer in which cells similar to those in the tumours were seen in the blood after death. Med J Aust 14:146–147

2. Salsbury AJ (1975) The significance of the circulating cancer cell. Cancer Treat Rev 2:55–72

3. Chang YS, Tomaso E di, McDonald DM, et al (2000) Mosaic blood vessels in tumors: frequency of cancer cells in contact with flowing blood. Proc Natl Acad Sci U S A 97:14608–14613

4. Massagué J and Obenauf AC (2016) Metastatic colonization by circulating tumour cells. Nature 529:298–306

5. Cristofanilli M, Budd GT, Ellis MJ, et al (2004) Circulating Tumor Cells, Disease Progression, and Survival in Metastatic Breast Cancer. N Engl J Med 351:781–791

6. Miller MC, Doyle GV, and Terstappen LWMM (2010) Significance of Circulating Tumor Cells Detected by the CellSearch System in Patients with Metastatic Breast Colorectal and Prostate Cancer. J Oncol 2010:617421

7. Aktas B, Kasimir-Bauer S, Heubner M, et al (2011) Molecular Profiling and Prognostic Relevance of Circulating Tumor Cells in the Blood of Ovarian Cancer Patients at Primary Diagnosis and After Platinum-Based Chemotherapy: Int J Gynecol Cancer 21:822–830

8. Minciacchi VR, Zijlstra A, Rubin MA, et al (2017) Extracellular vesicles for liquid biopsy in prostate cancer: where are we and where are we headed? Prostate Cancer Prostatic Dis 20:251–258

9. Thiele J-A, Bethel K, Králíčková M, et al (2017) Circulating Tumor Cells: Fluid Surrogates of Solid Tumors. Annu Rev Pathol Mech Dis 12:419–447

10. Hesari A, Moghadam SAG, Siasi A, et al (2017) Tumor-derived exosomes: potential biomarker or therapeutic target in breast cancer?: exosomes in breast cancer. J Cell Biochem

11. Alix-Panabieres C and Pantel K (2016) Clinical Applications of Circulating Tumor Cells and Circulating Tumor DNA as Liquid Biopsy. Cancer Discov

12. Cho WCS (2014) Emerging techniques in molecular detection of circulating tumor cells. Expert Rev Mol Diagn 14:131–134

13. Alix-Panabières C (2012) EPISPOT assay: detection of viable DTCs/CTCs in solid tumor patients. Recent Results Cancer Res Fortschritte Krebsforsch Prog Dans Rech Sur Cancer 195:69–76

14. Tewes M, Kasimir-Bauer S, Welt A, et al (2015) Detection of disseminated tumor cells in bone marrow and circulating tumor cells in blood of patients with early-stage male breast cancer. J Cancer Res Clin Oncol 141:87–92

15. Nagrath S, Sequist LV, Maheswaran S, et al (2007) Isolation of rare circulating tumour cells in cancer patients by microchip technology. Nature 450:1235–1239

16. Marrinucci D, Bethel K, Kolatkar A, et al (2012) Fluid biopsy in patients with metastatic prostate, pancreatic and breast cancers. Phys Biol 9:016003

17. Carlsson A, Nair VS, Luttgen MS, et al (2014) Circulating Tumor Microemboli Diagnostics for Patients with Non–Small-Cell Lung Cancer. J Thorac Oncol 9:1111–1119

18. Newton PK, Mason J, Bethel K, et al (2013) Spreaders and sponges define metastasis in lung cancer: a Markov chain Monte Carlo mathematical model. Cancer Res 73:2760–2769

19. Malihi PD, Morikado M, Welter L, et al (2018) Clonal diversity revealed by morphoproteomic and copy number profiles of single prostate cancer cells at diagnosis. Converg Sci Phys Oncol 4:015003

20. Ruiz C, Li J, Luttgen MS, et al (2015) Limited genomic heterogeneity of circulating melanoma cells in advanced stage patients. Phys Biol 12:016008

21. Bethel K, Luttgen MS, Damani S, et al (2014) Fluid phase biopsy for detection and characterization of circulating endothelial cells in myocardial infarction. Phys Biol 11:016002

22. Xuan J, Yu Y, Qing T, et al (2013) Next-generation sequencing in the clinic: Promises and challenges. Cancer Lett 340:284–295

23. Ranek L (1976) Cytophotometric studies of the DNA, nucleic acid and protein content of human liver cell nuclei. Acta Cytol 20:151–157

24. Baslan T, Kendall J, Rodgers L, et al (2012) Genome-wide copy number analysis of single cells. Nat Protoc 7:1024–1041

25. Dago AE, Stepansky A, Carlsson A, et al (2014) Rapid Phenotypic and Genomic Change in Response to Therapeutic Pressure in Prostate Cancer Inferred by High Content Analysis of Single Circulating Tumor Cells. PLoS ONE 9:e101777

26. Ni X, Zhuo M, Su Z, et al (2013) Reproducible copy number variation patterns among single circulating tumor cells of lung cancer patients. Proc Natl Acad Sci U S A 110:21083–21088

27. Navin N and Hicks J (2011) Future medical applications of single-cell sequencing in cancer. Genome Med 3:31

28. Navin N, Kendall J, Troge J, et al (2011) Tumour evolution inferred by single-cell sequencing. Nature 472:90–94

29. Baslan T, Kendall J, Ward B, et al (2015) Optimizing sparse sequencing of single cells for highly multiplex copy number profiling. Genome Res 25:714–724

30. Venkatraman ES and Olshen AB (2007) A faster circular binary segmentation algorithm for the analysis of array CGH data. Bioinformatics 23:657–663

31. Krzywinski M, Schein J, Birol I, et al (2009) Circos: An information aesthetic for comparative genomics. Genome Res 19:1639–1645

IV. Publication of a side project during PhD

Additionally to research regarding CTCs in CRC, a side project has been executed addressing the analysis of long non-coding RNA (lncRNA) expression in CRC and non-malignant tissue:

lncRNAs in healthy tissue have prognostic value in colorectal cancer

Thiele J-A, Hosek P, Kralovcova E, Ostasov P, Liska V, Bruha J, Vycital O, Rosendorf J, Kralickova M, Pitule P, lncRNAs in healthy tissue have prognostic value in colorectal cancer, Submitted to BMC Cancer: June 2018 **(IF_{2017}=3.28)**

lncRNAs in healthy tissue have prognostic value in colorectal cancer

J.-A. Thiele[1], P. Hosek[1], E. Kralovcova[1], P. Ostasov[1], V. Liska[2], J. Bruha[2], O. Vycital[2], J. Rosendorf[2], M. Kralickova[1,4], P. Pitule[1]

1 Biomedical Center, Faculty of Medicine in Pilsen, Charles University, 323 00 Pilsen, Czech Republic

2 Department of Surgery, Faculty of Medicine and University Hospital in Pilsen, Charles University, Czech Republic

3 Department of Oncology and Radiotherapeutics, Faculty of Medicine and University Hospital in Pilsen, Charles University, Czech Republic

4 Department of Histology and Embryology, Faculty of Medicine in Pilsen, Charles University, 301 00 Pilsen, Czech Republic

Corresponding Author:

Pavel Pitule: Biomedical Center, Faculty of Medicine in Pilsen, Charles University, 323 00 Pilsen, Czech Republic, e-mail: Pavel.Pitule@lfp.cuni.cz

Other Authors:

Jana-Aletta Thiele: Biomedical Center, Faculty of Medicine in Pilsen, Charles University, 323 00 Pilsen, Czech Republic, e-mail: Jana.A.Thiele@lfp.cuni.cz

Petr Hosek: Biomedical Center, Faculty of Medicine in Pilsen, Charles University, 323 00 Pilsen, Czech Republic, e-mail: petr.hosek@lfp.cuni.cz

Eva Kralovcova: Biomedical Center, Faculty of Medicine in Pilsen, Charles University, 323 00 Pilsen, Czech Republic, e-mail: Eva.Kralovcova@lfp.cuni.cz

Pavel Ostasov: Biomedical Center, Faculty of Medicine in Pilsen, Charles University, 323 00 Pilsen, Czech Republic, e-mail: Ostašov Pavel <Pavel.Ostasov@lfp.cuni.cz>

Vaclav Liska: University Hospital in Pilsen, Czech Republic, e-mail: LISKAV@fnplzen.cz

Jan Bruha: University Hospital in Pilsen, Czech Republic, e-mail: jan_bruha@volny.cz

Ondrej Vycital: University Hospital in Pilsen, Czech Republic, e-mail: 'vycitalo@fnplzen.cz'

Jachym Rosendorf: University Hospital in Pilsen, Czech Republic, e-mail: rjachym@seznam.cz

Milena Kralickova: University Hospital in Pilsen, Czech Republic, e-mail: Milena.Kralickova@lfp.cuni.cz

Pavel Pitule: University Hospital in Pilsen, Czech Republic, e-mail: Pavel.Pitule@lfp.cuni.cz

Abstract

Background: Colorectal cancer (CRC) is the third most frequent cause of cancer related death in Europe and clinically relevant biomarkers for therapy guidance and prediction of survival time are still rare. Long non-coding RNAs (lncRNAs) are a group of

RNAs that are over 200 nucleotides long and not translated into proteins, but with the ability to influence biological processes. There is evidence for their involvement in cancer as oncogenes, tumor suppressors or as regulators of cell proliferation and metastasis development.

Methods: In this retrospective study we used a quantitative PCR approach to measure the expression of nine lncRNAs previously shown to be involved in cancer progression – ANRIL, CCAT1, GUS5, linc-ROR, MALAT1, MIR155HG, PCAT1, SPRY4-IT1 and TUG1, in paired healthy and tumor tissue samples of CRC patients. Association of expression and expression ratios with survival and clinical characteristics was assessed by statistical analysis using univariable Cox proportional hazards model, Kaplan-Meier estimation with Gehan-Wilcoxon test, Mann-Whitney U test, Kruskal-Wallis ANOVA and Spearman's correlation.

Results: Comparison of expression in tumor tissue (TT) to healthy tissue (HT) showed significant upregulation of CCAT1 and linc-ROR in TT ($p<0.001$), whereas ANRIL, MIR155HG and MALAT1 were downregulated in TT ($p=0.001$, $p=0.010$, $p=0.001$, respectively). Linc-ROR was significantly associated with presence of synchronous metastases ($p=0.036$). For individual tissue types, lower MIR155HG expression in TT was correlated with shorter overall survival ($p=0.008$), and with shorter disease free survival ($p=0.040$). Within healthy tissue, expression ratios of CCAT1/ANRIL and CCAT1/MIR155HG revealed association with overall survival ($p=0.005$ and $p=0.006$). LncRNA expression ratios of selected lncRNAs in TT provided no information about patient's survival.

Conclusions: Our results confirmed that changes of expression of lncRNA between HT and TT carry a high potential for their use as prognostic biomarkers in CRC patients.

Moreover, ratios of CCAT1 with ANRIL and MIR155HG in HT also show potential as prognostic markers of high practical value as it might allow assessment without direct tumor sampling. This result also indicates that cancer progression is associated with detrimental system-wide changes in patient tissue, which might govern patient survival even after successful elimination of tumor or cancerous cells.

Keywords: Colorectal carcinoma, lncRNA, lncRNA ratio, MIR155HG, CCAT1, PCAT1

Background

Colorectal cancer (CRC) was in the year 2017 predicted to be the second leading cause of cancer-related death in men and the third leading cause for women in Europe [1]. As early stages of CRC are usually asymptomatic and biomarkers for its early detection are missing, less than 40% of patients are diagnosed with localized disease, which has a severe impact on the patient´s 5 year survival rate – it drops from 90% for localized disease down to 12% for patients with distant metastasis [2–4]. Standard treatment options for CRC consist mainly of radiation therapy or surgical intervention. Unfortunately, for 80-90% of patients with liver metastases resection is impossible [5]. This severe threat makes it more pressing to discover new biomarkers and potential therapeutic targets to predict disease evolution and uncover the relationships of epigenetic regulation of gene expression promoting metastasis.

After being brushed aside as ‘noise’ in the past, long non-coding RNAs (lncRNAs) have now been rediscovered as regulators of gene expression and are often abnormally expressed in tumor tissue [6]. In many studies from the last few years researchers try to uncover functionalities of lncRNAs to elucidate how they regulate gene expression, apoptosis or proliferation and to determine their role in the metastatic potential of cancer cells [7–13].

Based on the results of many of these studies we chose a set of nine lncRNAs for a study on Czech CRC patients across all American Joint Committee on Cancer (AJCC) stages (references and functions for each selected lncRNA are stated in **additional file 1**). The lncRNAs were chosen according to their potential to promote or influence CRC progression in general (Taurine Up-Regulated 1 (TUG1), Growth arrest-specific 5 (GAS5) and Metastasis Associated Lung Adenocarcinoma Transcript 1 (MALAT1)) [10,11,14,15] or through cell proliferation (sprouty homolog 4 intronic transcript 1 (SPRY4-IT1)), GAS5 and Colon Cancer Associated Transcript 1 (CCAT-1)) [10,12,16], cell migration (CDKN2B antisense RNA 1 (ANRIL)) [17], apoptosis (Prostate Cancer Associated Transcript 1 (PCAT1)) [8], being a regulator for known tumor suppressors like *p53* (Long Intergenic Non-Protein Coding Regulator of Reprogramming (linc-ROR)) [18] or inhibiting epithelial-mesenchymal transition (EMT) in other cancer types (MIR155 Host Gene (MIR155HG)) [19].

In oncological studies addressing lncRNAs, the common approach is to analyze differential expression between tumor and healthy tissue [8,9,20]. However, the use of normalized gene expression levels within individual tissue types has also been proven helpful in detecting metastatic risk in CRC [21], but not explored yet using lncRNAs. Expression ratios of two non-reference genes are also already applied in clinical practice for many cancers. For example, the diagnosis of mantle cell lymphoma through measuring the ratio of cyclin D1 and cyclin D3 (CCND1/CCND3) in tissue samples, peripheral blood or bone marrow [22]. However, to our knowledge it was again not applied to lncRNAs.

The aim of our retrospective study is to correlate expression levels of the selected lncRNAs (ANRIL, linc-ROR, CCAT1, PCAT1, SPRY4-IT1, TUG1, GAS5, MALAT1 and MIR155HG) measured by q-RT-PCR with survival and pathological features of CRC patients. To explore the potential of lncRNAs as biomarkers, we analyzed their differential expression in tumor tissue (TT) in

respect to healthy tissue (HT), their normalized expression in individual tissue types as well as expression ratios of pairs of lncRNAs in HT and TT separately.

Methods

Patient Cohort:

In this retrospective study we analyzed 63 adult patients from the Czech Republic with confirmed colorectal cancer operated between September 2012 and February 2014 at University hospital in Pilsen. All patients enrolled in the study agreed to the processing of their samples by signing informed consent and the study protocol was approved by the ethics committee of the Faculty of Medicine and University Hospital in Pilsen and complies with the International Ethical Guidelines for Biomedical Research Involving Human Subjects. Matching tissue samples from tumor and macroscopically healthy mucosa from the resected part of the colon (sampled as distant from the tumor as possible) were collected at the University Hospital in Pilsen during colorectal tumor resections. Anonymized clinical data were retrieved retrospectively from the hospital information system. An overview of characteristics of this cohort is presented in **Table 1**.

Table 1 | **Descriptive statistics for the analyzed patient cohort of colorectal cancer patients**

Characteristic	Category	No.	%
Gender	F	24	38.1
	M	39	61.9
Age (in years)	30 - 50	8	12.7
	50 - 70	32	50.8
	>70	23	36.5
T Stage	T1	1	1.6
	T2	14	22.2
	T3	40	63.5
	T4	6	9.5

	Unknown	2	3.2
N Stage	N0	39	61.9
	N1	13	20.6
	N2	10	15.9
	Unknown	1	1.6
M Stage	0	47	74.6
	1	11	17.5
	Unknown	5	7.9
Tumor Grade	G1	12	19.0
	G2	42	66.7
	G3	5	7.9
	Unknown	4	6.3
AJCC Staging	I	11	17.5
	II	19	30.2
	III	14	22.2
	IV	11	17.5
	Unknown	8	12.7

RNA extraction:

Frozen tissue samples of tumor or healthy mucosa were ground in liquid nitrogen and transferred into 1 ml of chilled TRI Reagent®RT (Molecular Research Center, Inc., Cincinnati, USA). RNA was then extracted following the manufacturer´s protocol and resuspended in molecular grade water. Isolated RNA was stored until further use in -80°C freezer.

RNA concentration was measured at the Infinite M200 (Tecan Trading AG, Männedorf, Switzerland) in the NanoQuant setting (260 nm absorbance) and purity determined by 230nm/260nm absorbance ratio. RNA integrity was tested by agarose gel electrophoresis. If degradation was detected, samples were not used for further analysis.

RT-PCR:

Reverse transcription PCR was performed using the RevertAid First Strand cDNA Synthesis Kit (ThermoFisher Scientific, #K1622) and 500 ng of total RNA in 20 µl reverse transcription reaction. Combined oligo(dT) and random hexamer primers were used, each in 2.5 µM final

concentration, to prime the reverse transcription to transcribe both mRNAs and lncRNAs. Reverse transcription and quality control PCR were performed in the T100 PCR system (Bio-Rad, ThermoFisher Scientific Inc., Waltham, USA). Quality and purity (absence of DNA contamination) of the cDNA was controlled by PCR using the PPP MasterMix (Top-Bio, s.r.o., Czech Republic) with the GAPDH primers from the RevertAid First Strand cDNA Synthesis Kit. Control PCR was ran for 40 cycles in 10 µl reactions using 5 µl of PPP MasterMix, 3.5 µl molecular water and 0.5 µl GAPDH primer mix with 1 µl of sample or reverse transcriptase negative control.

Quantitative-PCR (qPCR):

For the evaluation of lncRNA expression, gene-specific TaqMan® Gene Expression probes from ThermoFisher (ThermoFisher Scientific Inc., Waltham, USA) were used. Tested targets were MIR155HG (assay ID Hs01374569_m1), TUG1 (Hs00215501_m1), GAS5 (Hs03464472_m1), ANRIL (Hs04259476_m1), MALAT1 (Hs01910177_s1), SPRY4-IT1 (Hs03865501_s1) and PCAT-1 (Hs04275836_s1). The probes for CCAT1 and LINC-ROR were ordered from GeneriBiotech (GENERI BIOTECH s.r.o, Hradec Králové, Czech Republic), order number 00491-14, assay ID hCCAT1_Q1 (reference sequence NR_108049.1, amplifying all transcript variants of CCAT1) and assay ID hLINC-ROR_Q2 (reference sequence NR_048536.1) for LINC-ROR. As reference genes we used GAPDH (assay ID Hs02758991_g1), ACTB (Hs01060665_g1) and GUSB (Hs00939627_m1).

Expression of lncRNAs was measured in 10 µl in 96 well plates in duplicates with the following reaction parameters: 2 min at 50°C holding, 10 min at 95°C holding, then 42 cycles of: 15 sec at 95°C and 1 min at 60°C. We used the AppliedBiosystems® (AppliedBiosystems Corp., Foster City, USA) 7500 Fast Real-Time PCR cycler together with TaqMan® Gene

Expression Master Mix (ThermoFisher Scientific Inc., Waltham, USA) in standard ramp speed. All sample cDNA has been diluted 1:75 using 0.1 µg/ml yeast tRNA (ThermoFisher Scientific Inc., Waltham, USA) in nuclease free water. Each reaction contained 4.5 µl of diluted sample, 5 µl of TaqMan® Gene Expression Master Mix and 0.5 µl of the particular probe. Data analysis and manual quality control of automated thresholding was performed in the 7500 Software.

Statistical analysis:

Duplicates of Ct values were averaged before further processing. Ct values of all three housekeeping genes were averaged and the result used as the final reference, i.e. Ct(ref). All subsequent analyses were performed with Ct values and their differences, in particular: (i) -ΔΔCt representing expression fold change (i.e. differential expression) in TT with respect to HT ($-\Delta\Delta Ct = [Ct_{HT}(lncRNA) - Ct_{HT}(ref)] - [Ct_{TT}(lncRNA) - Ct_{TT}(ref)]$); (ii) -ΔCt representing normalized expression in a single tissue type ($-\Delta Ct = Ct(ref) - Ct(lncRNA)$); and (iii) negative value of the difference of lncRNA Ct values representing their ratio. The results of the analyses were converted into fold changes, normalized expressions and expression ratios as the last step before presentation while assuming 100% PCR efficiency (fold change = $2^{-\Delta\Delta Ct}$; normalized expression = $2^{-\Delta Ct}$; (expression A)/(expression B) = $2^{Ct(B)-Ct(A)}$).

Standard frequency tables and descriptive statistics were used to characterize the patient cohort. Significance of up- or down-regulation of the lncRNAs in TT was assessed by testing the -ΔΔCt values against zero location with Wilcoxon signed-rank test. Associations between expression descriptors and other clinical characteristics were analyzed using Mann-Whitney U test, Kruskal-Wallis ANOVA and Spearman's correlation.

For the purpose of survival analysis, disease-free survival (DFS) was determined from the date of surgery to the date of disease recurrence or death. The date of recurrence was set to the average date between the last negative and the first positive examination if the interval between the examinations was 180 days or less. In cases of longer examination interval, the recurrence date was set 90 days before the first positive examination. Overall survival (OS) was determined from the date of surgery to the date of death. Median follow-up was determined using the inverse Kaplan-Meier method. The significance of associations between lncRNA expression descriptors and survival times was assessed using univariable Cox proportional hazards model. In order to visualize these associations with Kaplan-Meier survival estimation plots, a threshold value needed to be set for each prognostic variable and the patients had to be stratified in two groups according to it. This threshold was found through automated optimization process implemented in Matlab (2014a, MathWorks Inc., Natick, MA, USA), in which the threshold value producing the smallest Log-rank p-value was determined and selected.

All reported p-values are two-tailed and the level of statistical significance was set at $\alpha = 0.05$. Statistical analysis was performed in Statistica (ver. 12 Cz, TIBCO Software Inc., Palo Alto, CA, USA).

Results

Patient characteristics

For our patient cohort the median overall survival and DFS has not been reached. The median follow-up time was 4.6 years. In our patient cohort we observed a three year OS of 85.1% and three-year DFS of 72.8%.

Expression fold change of lncRNAs in tumor tissue compared to healthy tissue

Quantitative PCR analysis showed differences in the lncRNA expression between the tested HT and TT samples (see **Figure 1**). Overexpression in TT is observed for the lncRNAs CCAT1 ($p<0.0001$) and linc-ROR ($p=0.0009$). Significant downregulation in tumor tissue is observed for the lncRNAs ANRIL ($p=0.0014$), MIR155HG ($p=0.0101$) and MALAT1 ($p=0.0006$).

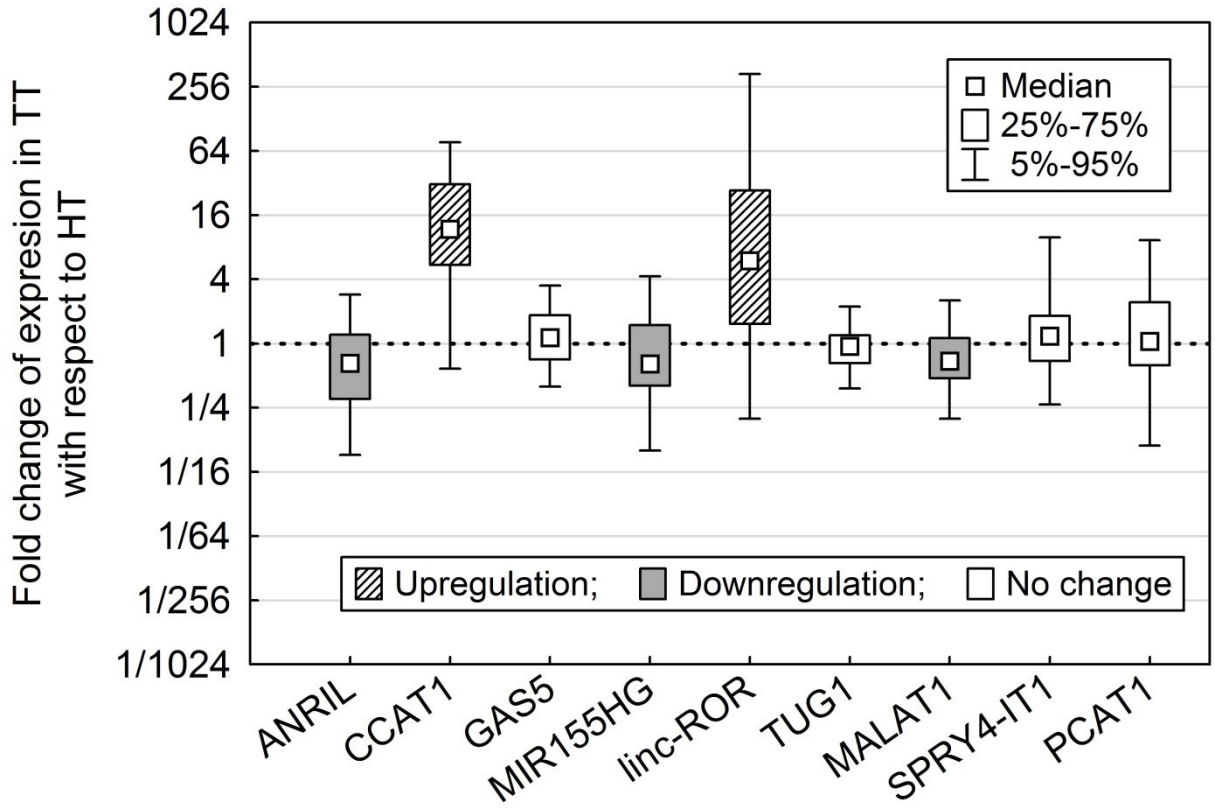

Figure 1 | **Expression of nine lncRNAs in tumor tissue compared to healthy tissue.** Boxplots of nine lncRNAs showing the fold change of their expression in tumor tissue vs. healthy tissue. Significantly up- or downregulated lncRNAs ($p < 0.05$) are marked with the respective patterns.

When investigating the expression of lncRNAs as a continuous predictor of DFS and OS using univariable Cox proportional hazards model, significant associations with DFS were observed for MIR155HG ($p=0.0402$) with a log2 hazard rate (HR) of 0.681 (the risk of disease

progression is decreased by 31.9% each time MIR155HG expression fold change is doubled) and for PCAT1 (p=0.0441) with a hazard rate of 0.731. Association with OS was only observed for MIR155HG (p=0.0079) with a hazard rate of 0.568).

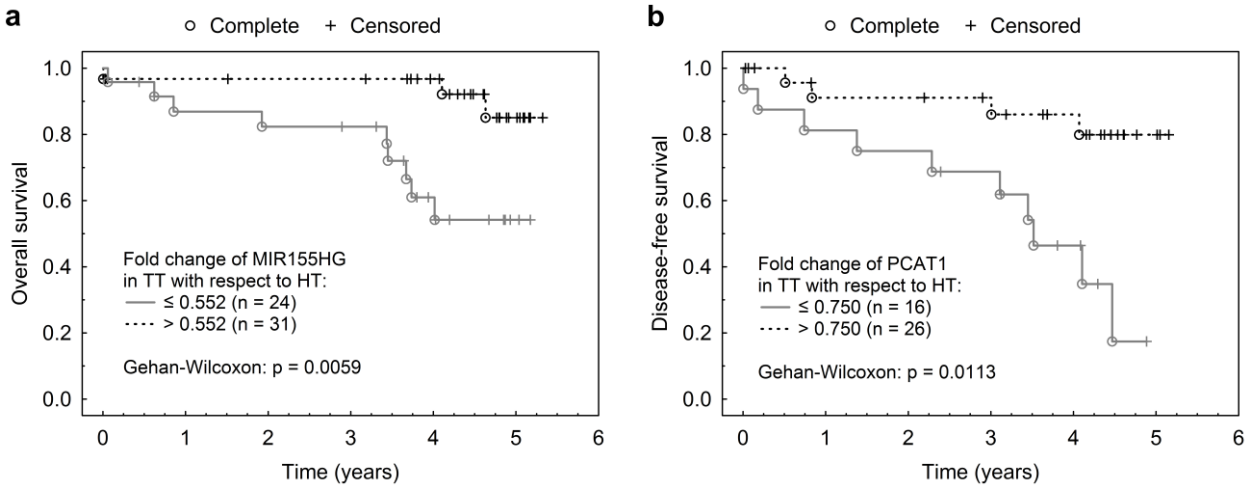

Figure 2 | **Association between MIR155HG and PCAT1 expression fold change and survival**. **a)** Kaplan Meier curve using optimized thresholds for fold change of MIR155HG and its association with OS and **b)** fold change of PCAT1 and its association with DFS

Subsequent comparison of Kaplan-Meier survival curves (**Fig. 2**) in patient groups based on lncRNA expression using an optimized threshold confirm the association between MIR155HG expression change and OS (p=0.0059, **Fig. 2a**) and PCAT1 expression change and DFS (p=0.0113, **Fig. 2b**).

When analyzing the associations of lncRNA expression changes with clinical characteristics (**Fig. 3**), we detected that the fold change of linc-ROR in TT compared to HT varies significantly between the four AJCC stages (Kruskal-Wallis: p=0.0126, **Fig. 3a**) and linc-ROR fold change is significantly higher in patients without distant metastases (Mann-Whitney U: p=0.0330, **Fig. 3b**). The expression change of MALAT1 varies significantly between T stages

(Kruskal-Wallis: p=0.0312, **Fig. 3c**) and is significantly different between T3 and T4 (p=0.0244, Multiple rank comparison). ANRIL fold change in TT is significantly higher in patients without lymph node involvement (Mann-Whitney U: p=0.0424, **Fig. 3d**). Also, an indication of a relationship between MIR155HG and M was observed, although without statistical significance (p=0.0612). Concerning other lncRNAs, Spearman correlation showed that with rising age TUG1 is more likely to be higher expressed in TT (p=0.0362, **additional file 2**).

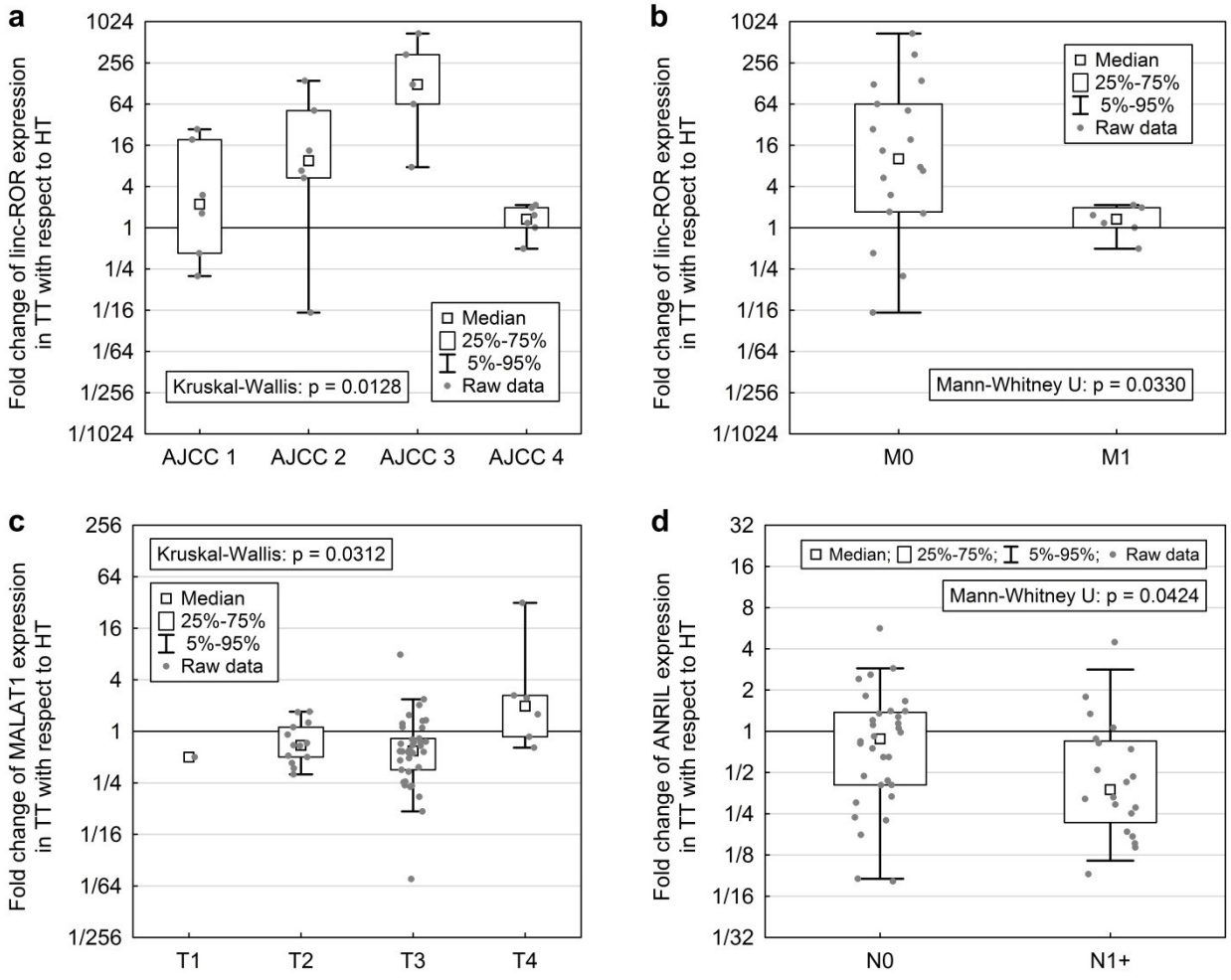

Figure 3 | Significant associations between expression fold change in TT vs. HT of the lncRNAs and tumor staging. Displayed is the ratio between healthy tissue (HT) and tumor tissue (TT). A value above 1 represents elevated expression in TT. **a)** The change in expression of linc-ROR differs significantly within the categories of AJCC staging and was largest in AJCC stage 3 samples. **b)** The expression change of linc-ROR is significantly

higher in patients without distant metastases. **c)** For MALAT1 the fold change varies significantly between the T stages. **d)** ANRIL expression change is significantly higher in patients without lymph node involvement (N0).

No correlation was detected between any lncRNA expression change and gender or cancer grade (G).

Expression of lncRNAs in healthy and tumor tissue

As a next step we analyzed the normalized expression of the nine tested lncRNAs in the two separate tissue types to test if their relative expression was associated to survival and/or clinical data.

Significant association for lncRNA expressions with OS has been detected using univariable Cox proportional hazard model for CCAT1 in HT (p=0.0329) with a HR of 0.634 (hazard drops by 36.6% every time CCAT1 expression in HT is doubled) and MIR155HG in TT (p=0.0464; HR=0.595). Patients with high expression of MIR155HG in TT are more likely to survive longer than those with low MIR155HG expression. Yet, patients with high CCAT1 expression in HT have a significant shorter OS. These associations have been visualized in Kaplan-Meier plots after applying an optimized threshold (**Fig. 4a and b**). The resulting survival curves confirm distinct differences in OS in relation to the expression of CCAT1 in HT (p=0.0017) and MIR155HG in TT (p=0.0154).

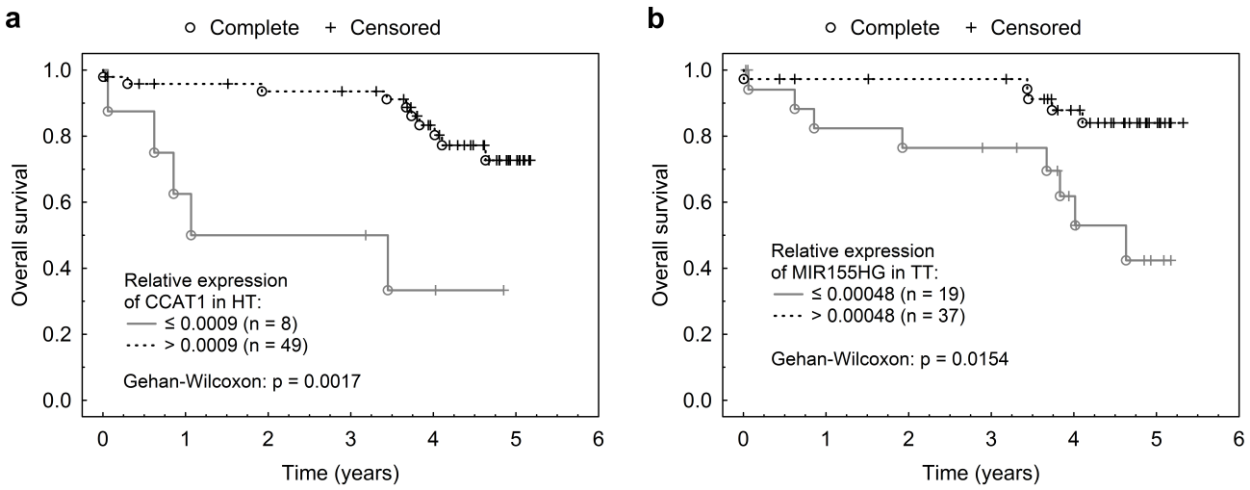

Figure 4 I **Associations of lncRNA normalized expressions in HT and TT with survival. a)** Kaplan Meier curve for association of CCAT1 expression in HT with OS and **b)** Kaplan Meier curve for MIR155HG expression in TT and its association with OS.

For lncRNA expressions in HT, linc-ROR shows significant differences in between the four AJCC stages (Kruskal-Wallis: p=0.0365, **Fig 5a**).

In TT MIR155HG is significantly more expressed in patients without metastases (M0) compared to those with metastases (Mann-Whitney U: p=0.0121, **Fig. 5b**). Expressions of linc-ROR and MALAT1 in TT are varying significantly within different tumor grades (Kruskal-Wallis: p=0.0348 and p=0.0454, **Fig. 5c and d**). Linc-ROR shows the highest expression in G3 tumors and MALAT1, which is down regulated in tumor tissue, is declining with tumor grade and is lowest in G3.

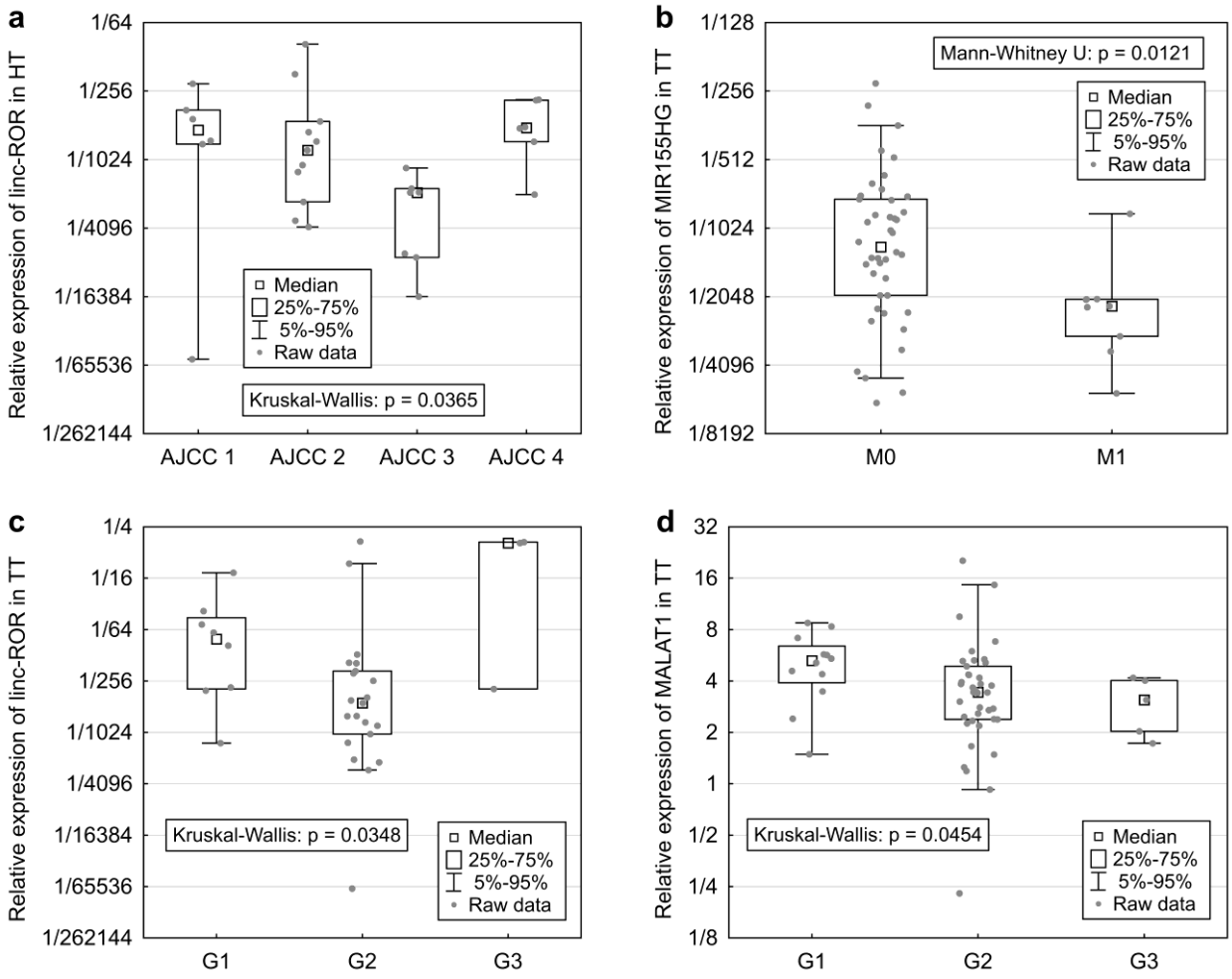

Figure 5 | Associations of lncRNA expressions in HT and TT with tumor characteristics. Displayed are the relative expression values in the respective tissue grouped by different clinical characteristics. A higher value corresponds to higher expression. **a)** The expression of linc-ROR in HT differs significantly within the categories of AJCC staging. **b)** The expression of MIR155HG in TT is significantly higher in patients without distant metastases. **c)** linc-ROR expression differs also significantly in TT between different tumor grades. **d)** MALAT1 expression in TT is on the edge of being associated to varying tumor grades

Expression ratios of two lncRNAs in healthy and tumor tissue

To analyze the power of combining multiple lncRNAs, we used the ratio of two expression values and calculated the p-values for all possible combinations in HT and TT using univariable Cox proportional hazard model. Due to the large-scale approach the threshold for significance was adjusted to $\alpha=0.01$ (i.e. $p<0.01$ was considered significant), but all p-values below 0.05 are discussed. After testing all combinations of lncRNA ratios, we found no association (neither $p<0.01$ nor $p<0.05$) with DFS or OS in TT.

For lncRNA expression ratios in HT, two ratios were significantly associated with OS, CCAT1/ANRIL ($p=0.0054$) with an HR of 0.646 (hazard drops by 35.4% every time the ratio is doubled) and CCAT1/MIR155HG ($p=0.0059$; HR=0.680). Four more ratios showed an indication of possible association with OS with a $p<0.05$ (**Fig. 6a**): CCAT1/GAS5 ($p=0.0197$; HR=0.672), MALAT1/ANRIL ($p=0.0243$; HR=0.641), SPRY4-IT1/ANRIL ($p=0.0386$; HR=0.665) and MALAT1/MIR155HG ($p=0.0486$; HR=0.651).

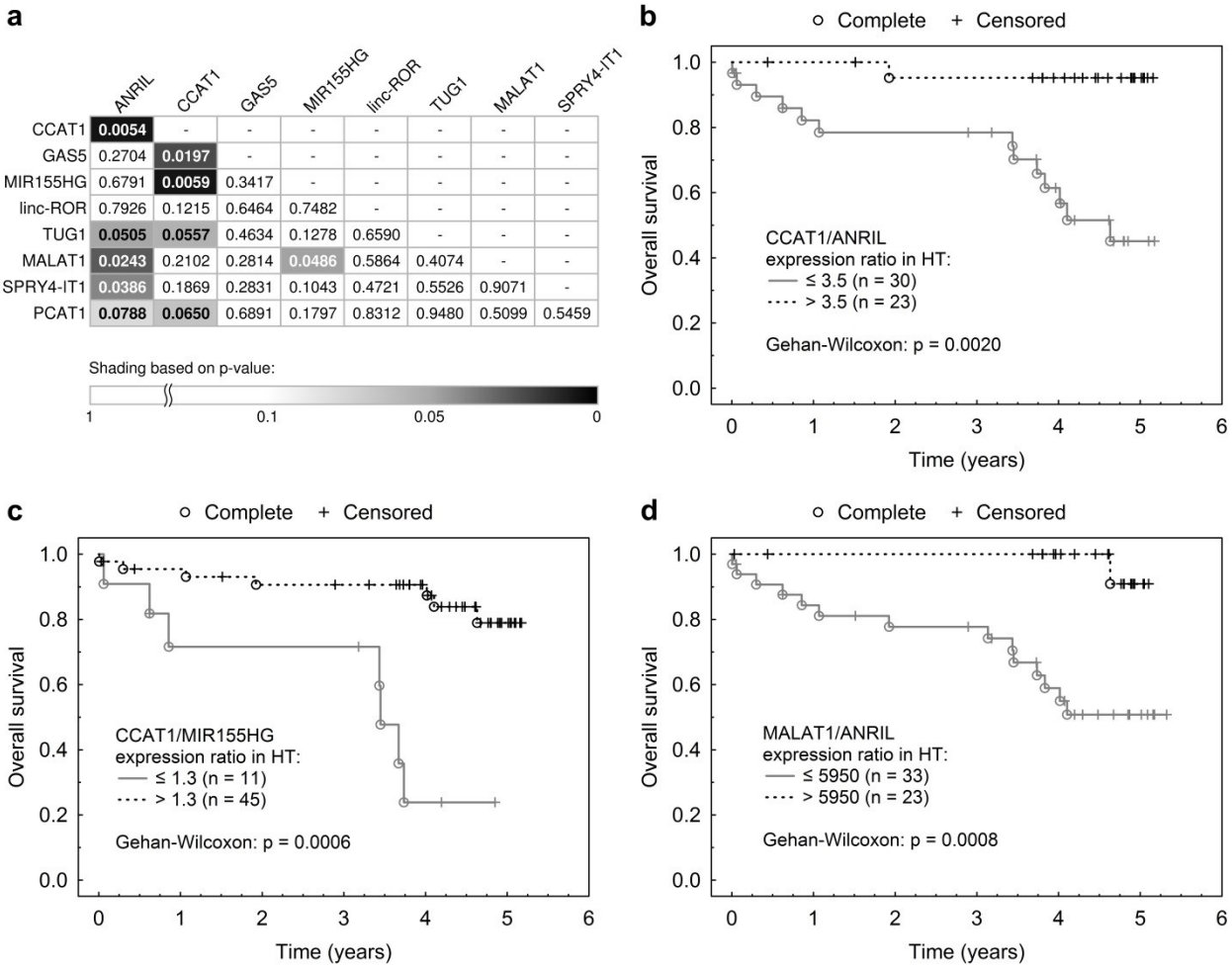

Figure 6 I **Associations of expression ratios of two lncRNAs in HT with OS. a)** Displayed are all possible expression ratios of the nine lncRNAs in a grate with their univariable Cox proportional hazard model p-values for association with OS. Each cell contains the p-value of the expression ratio of lncRNAs stated in the appropriate row and column headings. The p-values are identical for reciprocal ratios, i.e. for X/Y and Y/X. **b) - d)** Kaplan-Meier curves with an applied optimized threshold **b)** expression ratio of CCAT1/ANRIL in HT association with OS, **c)** CCAT1/MIR155HG in HT with OS and **d)** MALAT1/ANRIL in HT with OS.

When an optimized threshold was applied for Kaplan-Meier curves, the most significant prognostic ratio for survival (CCAT1/ANRIL) shows weak association with OS (Gehan-Wilcoxon: p=0.0020, **Fig. 6b**), but the ratios CCAT1/MIR155HG and MALAT1/ANRIL show a stronger association with OS (Gehan-Wilcoxon: p<0.001, **Fig.6c and d**). More Kaplan-Meier curves for lncRNA ratios can be found in **additional file 3**.

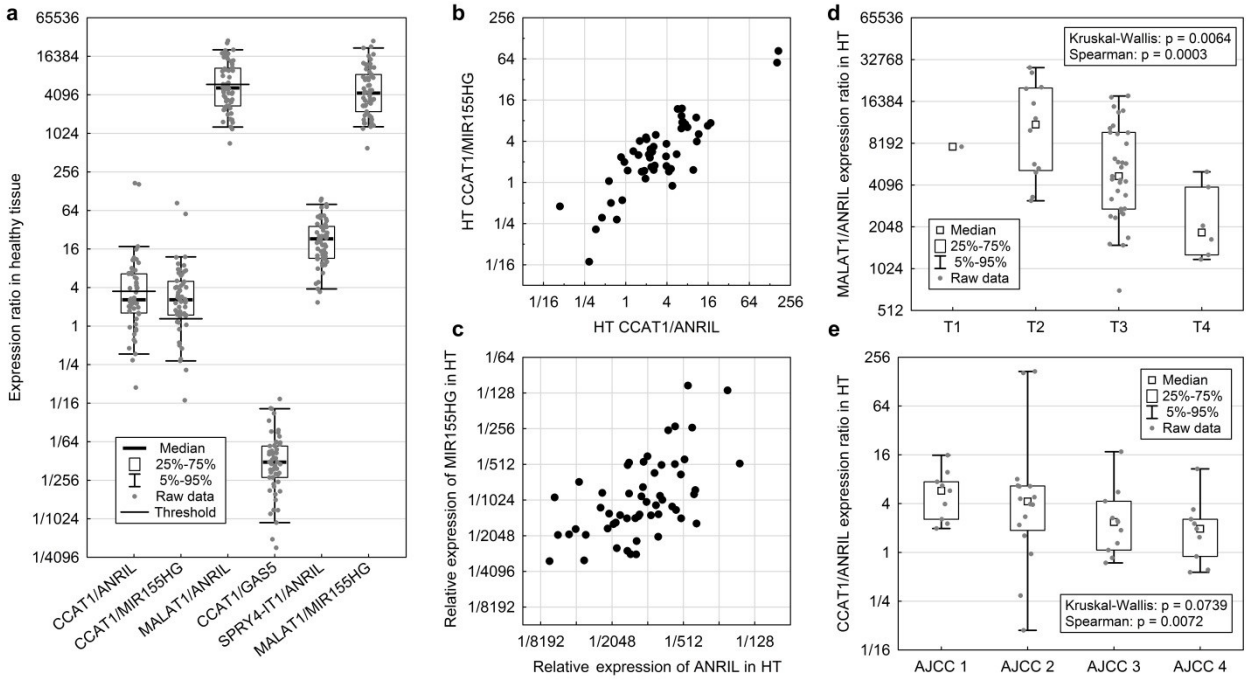

Figure 7I **Expression ratios of two lncRNAs in HT. a)** Displayed are the six lncRNA expression ratios in HT that are significantly or potentially associated with OS. For the ratios featured in Fig. 6 the values of the optimized threshold used for the Kaplan-Meier curves are also shown. **b)** Scatter plot for the correlation of the two most prognostic ratios in HT that are associated with OS. **c)** Scatter plot for the two relative expression values of lncRNAs in HT involved in the two strongest ratios except CCAT1. **d)** The ratio of MALAT1/ANRIL is associated with tumor stage and is decreasing with rising tumor staging. **e)** The expression ratio CCAT1/ANRIL is decreasing significantly with higher AJCC stage (Spearman: p=0.0072).

All ratios of HT lncRNA expression that are associated with OS are displayed in **Fig. 7a**. The two ratios with the highest association to OS show a strong mutual correlation (CCAT1/MIR155HG and CCAT1/ANRIL, p<0.001, **Fig. 7b**). This is due to two factors – first, CCAT1 being involved in both ratios, and second, relative expressions of the two differing lncRNAs being also moderately correlated (**Fig. 7c).** This indicates that both ratios might redundantly express the same prognostic feature.

The six most prognostic expression ratios in HT (two that showed associations with survival and four on the edge of significance) have been tested for correlations with clinical information. Association with T stage was detected for MALAT1/ANRIL (p=0.0064, **Fig. 7d**) with a strong trend to decrease with rising tumor stage (Spearman: p=0.0003).

Four expression ratios showed an association with AJCC stage of the tumor. The strongest association with AJCC stage showed the ratio CCAT1/ANRIL with the trend to decrease with rising AJCC stage (p=0.0072, **Fig. 7e**), second strongest has CCAT1/MIR155HG (p=0.0218) followed by MALAT1/ANRIL (p=0.0239) and SPRY4-IT1/ANRIL (p=0.0337). Displayed in **Figure 7** are just the strongest associations with clinical characteristics. Further data and figures for associations of lncRNA ratios of HT with clinical characteristics can be found in **additional file 3**.

No association was found for any of the tested ratios with G, N, M and age.

Discussion

Advanced colorectal cancer represents an often lethal disease, especially when metastasis has occurred – in that case, 5-year survival rate drops to 12% and in many cases resection turns out to be impossible. Thus the main goal in CRC research right now, apart from invention of new treatments, is early diagnosis and discovery of new predictors to allow to use already existing treatments efficiently or to allow precise prognosis for the development of the disease for improvement of patient's quality of life. Given the emerging role of lncRNAs as important epigenetic regulators of tumor development and disease progression in multiple cancers, we aimed to explore the full potential of lncRNA expression and discovered several significant and potentially useful associations.

Differential expression values in TT compared to HT

In our study we confirmed that one of the best known lncRNAs in CRC, CCAT1 has been upregulated in TT, similarly to previously published results [16]. CCAT1, also known as CARLo-5, is a lncRNA of the human gene desert region 8q24.21, that contains enhancers and lncRNAs, that have shown relevance in colorectal cancer and is known to influence cell growth and invasiveness in CRC [16,20]. However, in our study its differential expression in TT was not associated with survival.

The strongest association with survival was shown by the lncRNA MIR155HG (MIR155 host gene). This lncRNA is processed into microRNA miR-155 [23] that has been identified as a negative regulator of the tumor protein 53 (TP53) and mismatch repair (MMR) genes [24,25]. It is also known to be upregulated in many cancers and to promote migration and invasion of CRC cells [26]. However, latest studies have shown that miR-155 may also work as a tumor suppressor. Kim et al. have shown that miR-155 loss promotes tumor growth and Liu et al. that overexpressed miR-155 leads to apoptosis and suppresses cell proliferation [27,28]. Our comparison of MIR-155HG expression in tissues showed a downregulation in TT and the lower expression in TT was associated with shorter OS and DFS. Considering that MIR155HG transcript is processed into miR-155, these findings support the observation that miR-155 might work as a tumor suppressor. On the other hand, high rate of processing of MIR155HG into miR-155 might lead to reduced number of full MIR155HG transcripts present in the tissue. In either case the lncRNA MIR155HG may be a promising biomarker for a longitudinal study with early stage patients, to investigate its potential to predict metastasis formation before it's clinically detectable. Moreover, as MIR155HG expression is an independent predictor of OS and, being not associated with age, it might be useful for decision making for palliative therapy or any therapy at all when coupled with other

predictors of patient survival like age or overall physical health to improve patient's quality of life.

Correlation analysis of MIR155HG fold change with the other lncRNAs indicated a co-regulation with ANRIL in TT ($p<0.001$, Spearman test). As stated above, MIR155HG could be processed into miR-155 and this conversion has been identified as a driver for EMT in glioma [19]. ANRIL has been identified in pancreatic cancer to influence the cadherin-switch through inhibiting ATM-E2F1 and thereby activating EMT [29] so their correlation might be indicative of EMT in the tumor and thus representative for the tumor's metastatic potential. This is in line with results of the two published articles about ANRIL in CRC patients where ANRIL upregulation in CRC TT was associated with shorter survival and lymph node metastasis [17,30]. In our cohort ANRIL also significantly varies in N stages of CRC patients, but its differential expression is decreased in patients with lymph node metastasis. Our patient cohort has a different ethnical background and higher average age than those in the previously published CRC studies. ANRIL is also an epigenetic regulator of the tumor suppressor CDKN2A/B and is therefore influencing cell proliferation [31]. However, Cunnington et al. observed that specific genetic variants may influence ANRIL expression and lower expression of ANRIL was associated with diseases like melanoma and cardiovascular disease [32].

Linc-ROR has been shown to act as a repressor of p53 in CRC and also as a 'sponge' for the tumor suppressor miR-145. MiR-145 downregulates OCT4 and SOX2 through mRNA binding and degradation [18,33,34]. Our results confirm the upregulation of linc-ROR in tumor tissue of CRC patients across all stages. Looking at the AJCC stages and metastasis status however, we observed an association of higher linc-ROR expression in patients without distant

metastases and lower AJCC stages, though we do have to report that there were only six patients in the AJCC4 / M1 group.

PCAT1 was found to be upregulated in CRC and associated with poor survival [35], but this was not observed in our cohort. Higher PCAT1 expression was associated to longer DFS. This observation disagrees with its described function to repress the tumor suppressor BRCA2 and being a sponge for microRNAs of the cell growth pathway [36,37], showing potentially additional functions of PCAT1 in CRC.

In our study we have not observed TUG1 to be expressed differently in TT or associated with survival or other clinical characteristics except age although according to existing studies it should be significantly upregulated in TT and cell line experiments showed that TUG1 promotes cell proliferation and colon cancer cell migration [38,39].

SPRY4-IT1 has been demonstrated to predict poor prognosis in CRC and promotes metastasis by enhancing cell proliferation, EMT gene expression and invasion [12,40]. Our differential expression values for SPRY4-IT1 in TT show an upregulation (the median is above 1), but this was not significant in our samples. Possibly a larger cohort size would enhance significance.

LncRNA GAS5 has been reported to be less expressed in TT and negatively correlated with OS and cell line experiments point out that high GAS5 expression reduces apoptosis and cell growth in CRC. [41]. In our study GAS5 was equally expressed in HT and TT and did not show any association with survival or tumor stage.

In our comparative analysis between tissues, we also detected additional results contradicting current literature. MALAT1 is known to be upregulated in many cancers and associated with poor survival and tumor growth [15,42]. In our study it is significantly, despite only minimally downregulated in TT compared to HT, but then regarding T-staging, it

is significantly upregulated in T4 stage, which supports its promotion of tumor growth. A challenge in our cohort is the small group of progressed patients with metastases or large tumors, whereas the stated MALAT1 studies are usually composed of mainly T4 stage with lymph node involvement [15]. Additionally, recent results are supporting this theory by observing that MALAT1 expression is lower in tissue samples with M0 stage or N0, but that high expression is associated with distant metastases in the lung or liver [43].

Normalized expression in TT and HT

In addition to comparison of healthy and tumor tissue, we focused our study also on the individual tissue types and compared expression levels of target lncRNAs within the healthy or tumor tissue. In case of healthy tissue, we cannot exclude the possibility of micrometastatic infiltration, but the samples were macroscopically healthy. Information from the healthy tissue can describe the fitness of the normal tissue and therefore may have an effect on the patient´s prognosis. The normalized expression values showed only weak associations with survival and clinical characteristics. As stated above, differential expression of CCAT1 was not associated with survival. However, increased expression of CCAT1 in HT alone (but not in TT) showed significant association with shorter OS. To our knowledge, this type of analysis of normalized lncRNA expression in HT has never been considered before and the unexpected result might open the door to a completely new approach to lncRNA expression assessment in cancer. Additionally, linc-ROR expression in HT was significantly different through the AJCC stages, and while linc-ROR expression in TT was rising from AJCC1 till AJCC3, it was decreasing in HT. In the tumor tissue only, lower MIR155HG expression was

weakly associated with shorter OS, which is a novel finding supporting the recent evidence of tumor suppressor role of MIRH155HG.

lncRNA expression ratios

Our results prove that lncRNAs can function as biomarkers especially in HT and that even some that had no significant association while comparing expression in tissues or by itself in HT are now part of a strongly predictive ratio, like CCAT1 or ANRIL in HT. For others, like MIR155HG, the association with survival is now enhanced when used in an expression ratio in HT. Especially the lncRNAs MIR155HG, CCAT1 and ANRIL are early stages markers of cancer development that are responsible for promoting proliferation and migration (CCAT1, ANRIL) or are known to be observed to subsequently lead to disease (MIR155HG).

Conclusions

LncRNAs are emerging markers for various tumor types and they are starting to be used alongside the protein-coding genes for disease prognosis and prediction of treatment effect. In addition to the well-established differential expression between tumor and healthy tissue, the potential of lncRNAs in HT only should be studied in more detail, as it may give us information about the fitness of the macroscopically healthy tissue that will impact the patient´s prognosis after the tumor removal.

The results of this study show that a large scale lncRNA study might reveal advances of multi-variable analysis where larger patient groups show various combinations of lncRNA up- or downregulation. The approach of using expression ratios instead of fold-change data has great advantages. First, it allows biomarker analysis by q-PCR without having to rely on the stability and validation of housekeeping genes expression which represents a common

challenge and risk for misinterpretation in qPCR experiments [44]. Second, if two genes are combined in an expression ratio that have opposite prognostic effects it will enhance the predictive power immensely without additional markers or effort needed.

A prospective study with CCAT1/ANRIL and CCAT/MIR155HG ratios in healthy tissue concerning their role in survival prediction should be conducted to confirm our results. The two main players MIR155HG and ANRIL should be further analyzed regarding their pathway interaction and their influence of tumorigenesis. For observation of patients after tumor resection of CRC, PCAT1 and MIR155HG could serve as predictive biomarker in the clinic for probable survival time and metastatic potential in CRC patients.

Closing, we do want to point out that our patient cohort was very limited, heterogeneous and the length of patient follow-up time was mediocre in relation to the OS and DFS times. Also our cohort is racially homogeneous and results might not transfer to other ethnical groups. In future studies we would also recommend to include follow up sampling, blood sample analysis (as an easy accessible source of healthy cells) and consideration of therapy.

List of Abbreviations

AJCC	American Joint Committee on Cancer
ANRIL	CDKN2B antisense RNA 1
CCAT1	Colon Cancer Associated Transcript 1
CCND	Cyclin D
CRC:	Colorectal Cancer
Ct:	Cycle Threshold

DFS:	Disease Free Survival
EMT	Epithelial to mesenchymal transition
GAS5	Growth arrest-specific 5
HR	hazard rate
HT:	Healthy tissue
Linc-ROR	Long Intergenic Non-Protein Coding Regulator of Reprogramming
lncRNA:	Long non-coding RNA
MALAT1	Metastasis Associated Lung Adenocarcinoma Transcript 1
MIR155HG	MIR155 Host Gene
MMR:	Mismatch repair
OS:	Overall Survival
PCAT1	Prostate Cancer Associated Transcript 1
q-PCR:	Quantitative PCR
RT-PCR:	Reverse Transcription PCR
SPRY4-IT1	Sprouty homolog 4 intronic transcript 1
TP53:	Tumor protein p53
TT:	Tumor tissue
TUG1	Taurine Up-Regulated 1

Declarations

Ethics approval and consent to participate

All procedures performed in this study involving human participants were in accordance with the ethical standards of the University Hospital research committee. The study was approved by the joint ethics committee of the University Hospital in Pilsen and Faculty of Medicine in Pilsen. Written informed consent was obtained from all participants.

Consent for publication

Not applicable.

Availability of data and materials

The dataset generated and analyzed during the current study are not publicly available due to the institution's policies but are available from the corresponding author on reasonable request.

Competing Interests

The authors declare that they have no competing interests.

Funding

This work has been supported by Charles University Research Centre program UNCE/MED/006 "University Center of Clinical and Experimental Liver Surgery", by the Charles University Research Fund (Progres Q39) and by the National Sustainability Program I (NPU I) Nr. LO1503 provided by the Ministry of Education Youth and Sports of the Czech Republic.

Authors' contributions

JAT was the main contributor in writing the manuscript, performed most of the qPCR analyses and contributed to data interpretation. PH performed all statistical analyses and contributed to data interpretation and writing. EK performed all RNA extractions, control PCRs and parts of the qPCR experiments. PO contributed in data interpretation and revised the manuscript. VL was the responsible surgeon for making the study possible and JB, OV, MS, OF and OS were the clinical team of oncologists and surgeons collecting the tissue samples during surgery and obtaining patients consent. MK was overseeing the study and PP contributed through ideas, data interpretation, guidance for writing and also revised the manuscript. All authors read and approved the final manuscript.

Acknowledgements

Not applicable

Footnotes

References

[1] M. Malvezzi, G. Carioli, P. Bertuccio, P. Boffetta, F. Levi, C. La Vecchia, E. Negri, European cancer mortality predictions for the year 2017, with focus on lung cancer, Ann. Oncol. 28 (2017) 1117–1123. doi:10.1093/annonc/mdx033.

[2] R.L. Siegel, K.D. Miller, S.A. Fedewa, D.J. Ahnen, R.G.S. Meester, A. Barzi, A. Jemal, Colorectal cancer statistics, 2017: Colorectal Cancer Statistics, 2017, CA. Cancer J. Clin. 67 (2017) 177–193. doi:10.3322/caac.21395.

[3] What Are the Survival Rates for Colorectal Cancer, by Stage?, (n.d.). https://www.cancer.org/cancer/colon-rectal-cancer/detection-diagnosis-staging/survival-rates.html (accessed February 2, 2018).

[4] Can Colorectal Polyps and Cancer Be Found Early?, (n.d.). https://www.cancer.org/cancer/colon-rectal-cancer/detection-diagnosis-staging/detection.html (accessed February 2, 2018).

[5] L.M. Pasetto, A. Jirillo, G. Iadicicco, E. Rossi, M.K. Paris, S. Monfardini, FOLFOX versus FOLFIRI: a comparison of regimens in the treatment of colorectal cancer metastases, Anticancer Res. 25 (2005) 563–576.

[6] X. Hu, A.K. Sood, C.V. Dang, L. Zhang, The role of long noncoding RNAs in cancer: the dark matter matters, Curr. Opin. Genet. Dev. 48 (2018) 8–15. doi:10.1016/j.gde.2017.10.004.

[7] J. Zhang, P. Zhang, L. Wang, H. -l. Piao, L. Ma, Long non-coding RNA HOTAIR in carcinogenesis and metastasis, Acta Biochim. Biophys. Sin. 46 (2014) 1–5. doi:10.1093/abbs/gmt117.

[8] L. Qiao, X. Liu, Y. Tang, Z. Zhao, J. Zhang, Y. Feng, Down regulation of the long non-coding RNA PCAT-1 induced growth arrest and apoptosis of colorectal cancer cells, Life Sci. (2017). doi:10.1016/j.lfs.2017.08.024.

[9] X. Shen, Y. Bai, B. Luo, X. Zhou, Upregulation of lncRNA BANCR associated with the lymph node metastasis and poor prognosis in colorectal cancer, Biol. Res. 50 (2017). doi:10.1186/s40659-017-0136-5.

[10] D. Yin, X. He, E. Zhang, R. Kong, W. De, Z. Zhang, Long noncoding RNA GAS5 affects cell proliferation and predicts a poor prognosis in patients with colorectal cancer, Med. Oncol. 31 (2014). doi:10.1007/s12032-014-0253-8.

[11] L. Wang, Z. Zhao, W. Feng, Z. Ye, W. Dai, C. Zhang, J. Peng, K. Wu, Long non-coding RNA TUG1 promotes colorectal cancer metastasis via EMT pathway, Oncotarget. 7 (2016) 51713–51719. doi:10.18632/oncotarget.10563.

[12] J. Jin, Z. Chu, P. Ma, Y. Meng, Y. Yang, Long non-coding RNA SPRY4-IT1 promotes proliferation and invasion by acting as a ceRNA of miR-101-3p in colorectal cancer cells, Tumor Biol. 39 (2017) 101042831771625. doi:10.1177/1010428317716250.

[13] P. Zhou, L. Sun, D. Liu, C. Liu, L. Sun, Long Non-Coding RNA lincRNA-ROR Promotes the Progression of Colon Cancer and Holds Prognostic Value by Associating with miR-145, Pathol. Oncol. Res. 22 (2016) 733–740. doi:10.1007/s12253-016-0061-x.

[14] X. Lu, Z. Liu, X. Ning, L. Huang, B. Jiang, The long noncoding RNA HOTAIR promotes colorectal cancer progression by sponging miR-197, Oncol. Res. Featur. Preclin. Clin. Cancer Ther. (2017). doi:10.3727/096504017X15105708598531.

[15] H.-T. Zheng, D.-B. Shi, Y.-W. Wang, X.-X. Li, Y. Xu, P. Tripathi, W.-L. Gu, G.-X. Cai, S.-J. Cai, High expression of lncRNA MALAT1 suggests a biomarker of poor prognosis in colorectal cancer, Int. J. Clin. Exp. Pathol. 7 (2014) 3174–3181.

[16] X. He, X. Tan, X. Wang, H. Jin, L. Liu, L. Ma, H. Yu, Z. Fan, C-Myc-activated long noncoding RNA CCAT1 promotes colon cancer cell proliferation and invasion, Tumor Biol. 35 (2014) 12181–12188. doi:10.1007/s13277-014-2526-4.

[17] Y. Sun, Z.-P. Zheng, H. Li, H.-Q. Zhang, F.-Q. Ma, ANRIL is associated with the survival rate of patients with colorectal cancer, and affects cell migration and invasion in vitro, Mol. Med. Rep. 14 (2016) 1714–1720. doi:10.3892/mmr.2016.5409.

[18] A. Zhang, N. Zhou, J. Huang, Q. Liu, K. Fukuda, D. Ma, Z. Lu, C. Bai, K. Watabe, Y.-Y. Mo, The human long non-coding RNA-RoR is a p53 repressor in response to DNA damage, Cell Res. 23 (2013) 340–350. doi:10.1038/cr.2012.164.

[19] X. Wu, Y. Wang, T. Yu, E. Nie, Q. Hu, W. Wu, T. Zhi, K. Jiang, X. Wang, X. Lu, H. Li, N. Liu, J. Zhang, Y. You, Blocking MIR155HG/miR-155 axis inhibits mesenchymal transition in glioma, Neuro-Oncol. 19 (2017) 1195–1205. doi:10.1093/neuonc/nox017.

[20] Q. Li, Y. Dai, F. Wang, S. Hou, Differentially expressed long non-coding RNAs and the prognostic potential in colorectal cancer, Neoplasma. 63 (2016) 977–983. doi:10.4149/neo_2016_617.

[21] L. Zhong, J. Liu, Y. Hu, W. Wang, F. Xu, W. Xu, J. Han, E. Biskup, STK31 as novel biomarker of metastatic potential and tumorigenicity of colorectal cancer, Oncotarget. 8 (2017). doi:10.18632/oncotarget.15396.

[22] C.D. Jones, K.H. Darnell, R.A. Warnke, J.L. Zehnder, CyclinD1/CyclinD3 Ratio by Real-Time PCR Improves Specificity for the Diagnosis of Mantle Cell Lymphoma, J. Mol. Diagn. 6 (2004) 84–89. doi:10.1016/S1525-1578(10)60494-1.

[23] P.S. Eis, W. Tam, L. Sun, A. Chadburn, Z. Li, M.F. Gomez, E. Lund, J.E. Dahlberg, Accumulation of miR-155 and BIC RNA in human B cell lymphomas, Proc. Natl. Acad. Sci. 102 (2005) 3627–3632. doi:10.1073/pnas.0500613102.

[24] N. Valeri, P. Gasparini, M. Fabbri, C. Braconi, A. Veronese, F. Lovat, B. Adair, I. Vannini, F. Fanini, A. Bottoni, S. Costinean, S.K. Sandhu, G.J. Nuovo, H. Alder, R. Gafa, F. Calore, M. Ferracin, G. Lanza, S. Volinia, M. Negrini, M.A. McIlhatton, D. Amadori, R. Fishel, C.M. Croce, Modulation of mismatch repair and genomic stability by miR-155, Proc. Natl. Acad. Sci. 107 (2010) 6982–6987. doi:10.1073/pnas.1002472107.

[25] G. Teng, F.N. Papavasiliou, Shhh! Silencing by microRNA-155, Philos. Trans. R. Soc. B Biol. Sci. 364 (2009) 631–637. doi:10.1098/rstb.2008.0209.

[26] G.-J. Zhang, H.-X. Xiao, H.-P. Tian, Z.-L. Liu, S.-S. Xia, T. Zhou, Upregulation of microRNA-155 promotes the migration and invasion of colorectal cancer cells through the regulation of claudin-1 expression, Int. J. Mol. Med. 31 (2013) 1375–1380. doi:10.3892/ijmm.2013.1348.

[27] J. Liu, Z. Chen, J. Xiang, X. Gu, MicroRNA-155 acts as a tumor suppressor in colorectal cancer by targeting CTHRC1 inï¿½vitro, Oncol. Lett. (2018). doi:10.3892/ol.2018.8069.

[28] S. Kim, J.H. Song, S. Kim, P. Qu, B.K. Martin, W.S. Sehareen, D.C. Haines, P.C. Lin, S.K. Sharan, S. Chang, Loss of oncogenic miR-155 in tumor cells promotes tumor growth by enhancing C/EBP-β-mediated MDSC infiltration, Oncotarget. 7 (2016). doi:10.18632/oncotarget.7150.

[29] S. Chen, J.-Q. Zhang, J.-Z. Chen, H.-X. Chen, F.-N. Qiu, M.-L. Yan, Y.-L. Chen, C.-H. Peng, Y.-F. Tian, Y.-D. Wang, The over expression of long non-coding RNA ANRIL promotes epithelial-mesenchymal transition by activating the ATM-E2F1 signaling pathway in pancreatic cancer: An in vivo and in vitro study, Int. J. Biol. Macromol. 102 (2017) 718–728. doi:10.1016/j.ijbiomac.2017.03.123.

[30] Z. Sun, C. Ou, W. Ren, X. Xie, X. Li, G. Li, Downregulation of long non-coding RNA ANRIL suppresses lymphangiogenesis and lymphatic metastasis in colorectal cancer, Oncotarget. 7 (2016). doi:10.18632/oncotarget.9868.

[31] A. Congrains, K. Kamide, M. Ohishi, H. Rakugi, ANRIL: Molecular Mechanisms and Implications in Human Health, Int. J. Mol. Sci. 14 (2013) 1278–1292. doi:10.3390/ijms14011278.

[32] M.S. Cunnington, M. Santibanez Koref, B.M. Mayosi, J. Burn, B. Keavney, Chromosome 9p21 SNPs Associated with Multiple Disease Phenotypes Correlate with ANRIL Expression, PLoS Genet. 6 (2010) e1000899. doi:10.1371/journal.pgen.1000899.

[33] P. Zhou, L. Sun, D. Liu, C. Liu, L. Sun, Long Non-Coding RNA lincRNA-ROR Promotes the Progression of Colon Cancer and Holds Prognostic Value by Associating with miR-145, Pathol. Oncol. Res. 22 (2016) 733–740. doi:10.1007/s12253-016-0061-x.

[34] N. Xu, T. Papagiannakopoulos, G. Pan, J.A. Thomson, K.S. Kosik, MicroRNA-145 Regulates OCT4, SOX2, and KLF4 and Represses Pluripotency in Human Embryonic Stem Cells, Cell. 137 (2009) 647–658. doi:10.1016/j.cell.2009.02.038.

[35] X. Ge, Y. Chen, X. Liao, D. Liu, F. Li, H. Ruan, W. Jia, Overexpression of long noncoding RNA PCAT-1 is a novel biomarker of poor prognosis in patients with colorectal cancer, Med. Oncol. 30 (2013). doi:10.1007/s12032-013-0588-6.

[36] J.R. Prensner, W. Chen, M.K. Iyer, Q. Cao, T. Ma, S. Han, A. Sahu, R. Malik, K. Wilder-Romans, N. Navone, C.J. Logothetis, J.C. Araujo, L.L. Pisters, A.K. Tewari, C.E. Canman, K.E. Knudsen, N. Kitabayashi, M.A. Rubin, F. Demichelis, T.S. Lawrence, A.M. Chinnaiyan, F.Y. Feng, *PCAT-1* , a Long Noncoding RNA, Regulates BRCA2 and Controls Homologous Recombination in Cancer, Cancer Res. 74 (2014) 1651–1660. doi:10.1158/0008-5472.CAN-13-3159.

[37] PCAT1 Gene - GeneCards | PCAT1 RNA Gene, (n.d.). http://www.genecards.org/cgi-bin/carddisp.pl?gene=PCAT1 (accessed April 9, 2018).

[38] H. Zhai, M. Sui, X. Yu, Z. Qu, J. Hu, H. Sun, H. Zheng, K. Zhou, L. Jiang, Overexpression of Long Non-Coding RNA TUG1 Promotes Colon Cancer Progression, Med. Sci. Monit. 22 (2016) 3281–3287. doi:10.12659/MSM.897072.

[39] J. Sun, C. Ding, Z. Yang, T. Liu, X. Zhang, C. Zhao, J. Wang, The long non-coding RNA TUG1 indicates a poor prognosis for colorectal cancer and promotes metastasis by affecting epithelial-mesenchymal transition, J. Transl. Med. 14 (2016). doi:10.1186/s12967-016-0786-z.

[40] F. Shen, W.-S. Cai, Z. Feng, J. Chen, J. Feng, Q. Liu, Y. Fang, K. Li, H. Xiao, J. Cao, B. Xu, Long non-coding RNA SPRY4-IT1 pormotes colorectal cancer metastasis by regulate epithelial-mesenchymal transition, Oncotarget. (2017). doi:10.18632/oncotarget.10407.

[41] Y. Yang, Z. Shen, Y. Yan, B. Wang, J. Zhang, C. Shen, T. Li, C. Ye, Z. Gao, G. Peng, Y. Ye, K. Jiang, S. Wang, Long non-coding RNA GAS5 inhibits cell proliferation, induces G0/G1 arrest and apoptosis, and functions as a prognostic marker in colorectal cancer, Oncol. Lett. 13 (2017) 3151–3158. doi:10.3892/ol.2017.5841.

[42] Q. Ji, L. Zhang, X. Liu, L. Zhou, W. Wang, Z. Han, H. Sui, Y. Tang, Y. Wang, N. Liu, J. Ren, F. Hou, Q. Li, Long non-coding RNA MALAT1 promotes tumour growth and metastasis in colorectal cancer through binding to SFPQ and releasing oncogene PTBP2 from SFPQ/PTBP2 complex, Br. J. Cancer. 111 (2014) 736–748. doi:10.1038/bjc.2014.383.

[43] Z.H. Kwok, V. Roche, X.H. Chew, A. Fadieieva, Y. Tay, A non-canonical tumor suppressive role for the long non-coding RNA MALAT1 in colon and breast cancers, Int. J. Cancer. (2018). doi:10.1002/ijc.31386.

[44] B. Kozera, M. Rapacz, Reference genes in real-time PCR, J. Appl. Genet. 54 (2013) 391–406. doi:10.1007/s13353-013-0173-x.

[45] B. Cai, X.Q. Song, J.P. Cai, S. Zhang, HOTAIR: a cancer-related long non-coding RNA, Neoplasma. 61 (2014) 379–391.

[46] T. Ozawa, T. Matsuyama, Y. Toiyama, N. Takahashi, T. Ishikawa, H. Uetake, Y. Yamada, M. Kusunoki, G. Calin, A. Goel, CCAT1 and CCAT2 long noncoding RNAs, located within the 8q.24.21 'gene desert', serve as important prognostic biomarkers in colorectal cancer, Ann. Oncol. (2017). doi:10.1093/annonc/mdx248.