Dissertation
submitted to the
Combined Faculties for the Natural Sciences and for Mathematics
of the Ruperto-Carola University of Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

presented by
Diplom-Biochemiker Thomas Niemietz
born in: Weimar
Oral-examination: .....

Referees:
First: **Prof. Dr. Philipp Beckhove**
Second: **Prof. Dr. Jürgen Weitz**

Phenotypic and molecular characterization of colorectal cancer-derived circulating and disseminated tumor cells

Table of contents

Abstract

Tumors shed large numbers of cells into the vasculature. It is supposed that these cells give rise to metastases that are the major cause of death from cancer. However, the isolation and characterization of the so called circulating tumor cells (CTCs) is still challenging since cells are rarely to find among millions of normal leukocytes. The objective of this thesis was to characterize colorectal cancer (CRC) - associated CTCs and disseminated tumor cells (DTCs) from bone marrow samples. We studied CTC incidence in blood samples of CRC patients with the help of the FDA-cleared CellSearch^{TM} system. We found that the number of patients with CTCs and the amount of CTCs was significantly correlated with the stage of disease. In addition, a significant higher rate of patients with CTCs in tumor-draining venous blood compared to the central venous blood was found. Furthermore, genomic analyses of single CTCs have been performed. This comprised the evaluation of typical CRC-associated mutations such as point mutations in *TP53*, *BRAF*, *KRAS* as well as the detection of microsatellite instability (MSI). Additionally, some CTCs were used to study global chromosomal aberrations by comparative genomic hybridization technique. Results revealed a remarkable genomic heterogeneity among the CTCs of single patients. Moreover, we detected several cases of genomic disparity among CTCs and the prevailing clone of matched cancer tissue. To evaluate differential gene expression that might enable tumor cell dissemination, we studied CTCs that were obtained from patients' blood samples and from an othotopic mouse model of metastatic CRC. Gene expression profiles of CTCs and cell samples from cancer tissue were compared. The down-regulation of cell adhesion molecules such as E-cadherin might be involved in tumor cell dissemination. However, a significant epithelial-mesenchymal transition (EMT) in CTCs could not be confirmed. DTCs isolated from human bone marrow samples seem to adopt a de-differentiated phenotype and lack expression of CK20 and CK19. Absence of EGFR and the proliferation marker Ki67 confirmed previous reports of dormancy in DTCs.

With the present work we were able to demonstrate that the molecular characterization of single CTCs is feasible. However, further studies are required to increase the knowledge about the molecular traits of CTCs which might help to improve therapy and prognosis in CRC.

Zusammenfassung

Krebszellen werden von Tumoren zahlreich ins Blut abgegeben. Man vermutet, dass diese Zellen Metastasen, die Hauptursache krebsbedingter Todesfälle, initiieren können. Jedoch ist die Isolation und Charakterisierung dieser so genannten zirkulierenden Tumorzellen (ZTZs) immer noch eine Herausforderung, da sie unter Millionen von Leukozyten nur schwer zu detektieren sind. Diese Arbeit wurde mit dem Ziel erstellt ZTZs sowie ins Knochenmark disseminierte Tumorzellen (DTZ) im kolorektalen Karzinom (KRK) zu charakterisieren. Das Vorkommen von ZTZs in Blutproben von KRK Patienten wurde mit dem CellSearch^{TM} System untersucht. Sowohl Anzahl der Patienten mit ZTZs als auch Menge von ZTZs korrelierten signifikant mit dem Krankheitsstadium. Es fanden sich signifikant mehr Patienten mit ZTZs in tumordrainierenden venösem Blut im Vergleich zu zentralvenösem Blut. Zusätzlich wurden genomische Einzelzellanalysen an ZTZs durchgeführt. Dies umfasste die Analyse von typischen KRK-assoziierten Punktmutationen in *TP53*, *KRAS* und *BRAF* sowie das Detektieren von Mikrosatelliteninstabilität. Einige ZTZs wurden für die Analyse von chromosomalen Aberrationen mittels CGH untersucht. Eine bemerkenswerte genomische Heterogenität innerhalb der ZTZ Populationen einzelner Patienten wurde gefunden. Außerdem traten häufig genomische Unterschiede zwischen ZTZs und dem dazugehörigen Krebsgewebe auf. Da differenzielle Genexpression vermutlich die Dissemination ermöglicht, wurden ZTZs, die aus Patientenblut und aus einem orthotopen Metastasierungsmodel in der Maus isoliert wurden, analysiert. Die Genexpressionsprofile der ZTZs und Zellproben des Krebsgewebes wurden verglichen. Die Runterregulation von Zelladhäsionsmolekülen wie E-cadherin könnte an der Tumorzelldissemination beteiligt sein. Allerdings konnte eine signifikante Epitheliale-Mesenchymale-Transition der ZTZs nicht bestätigt werden. DTZs, die aus menschlichen Knochenmarkproben isoliert wurden, scheinen einen dedifferenzierten Phänotyp anzunehmen. Die Differenzierungsmarker CK20 und CK19 wurden nicht detektiert. Berichte über eine niedrige Expression von Ki67 und EGFR, die auf eine proliferative Quieszenz in DTZs hindeuten, ließen sich hingegen bestätigen. Die vorliegende Arbeit zeigt, dass eine umfassende molekulare Charakterisierung von ZTZs und DTZs möglich ist. Um mehr über die molekularen Eigenschaften dieser Zellen zu erfahren sind jedoch weitere Studien erforderlich. Diese könnten die Therapie und Prognose des KRK verbessern helfen.

Aim

The objectives of this thesis were the detection, isolation and characterization of colorectal cancer (CRC) - derived circulating and disseminated tumor cells (CTCs /DTCs)). To increase our knowledge about the early steps of metastasis we were focused on the genomic and transcriptional profiles of CTCs and DTCs. Due to the lack of reliable and standardized techniques to isolate and analyze CTCs obtained from human blood samples, an efficient CTC isolation protocol had to be designed. Transcriptional profiles of solid tumor samples and CTCs had to be ascertained to analyze disparities of the mRNA expression profile during and after the dissemination process. Additionally, an orthotopic mouse model was established to mimic cancer cell dissemination. CTCs that were obtained from the mouse model ought to provide further information about the molecular mechanisms involved in tumor cell dissemination.

Introduction

Colorectal cancer

CRC is ranked as the third most common cancer worldwide [1]. In Europe about 250.000 new cases of colon cancer are diagnosed each year [2]. However, the majority of the patients does not succumb to the primary tumor but dies due to metastases with a five-year survival rate of only 30-40% [3]. Surgery is the primary treatment for CRC; nevertheless, CRC seems to be curable only when diagnosed and treated at an early stage of disease [4, 5].

CRC progression is mostly staged by the tumor-node-metastasis (TNM) classification. This system is based on the depth of invasion of the bowel wall, the extent of lymph node involvement, and the presence of distant metastasis. Beyond that, an additional system for overall staging was suggested by the Union International contre le Cancer (UICC). It is referred as Roman Numeral staging that classifies increased tumor progression with roman numerals from 0 – IV (Figure 1) [6].

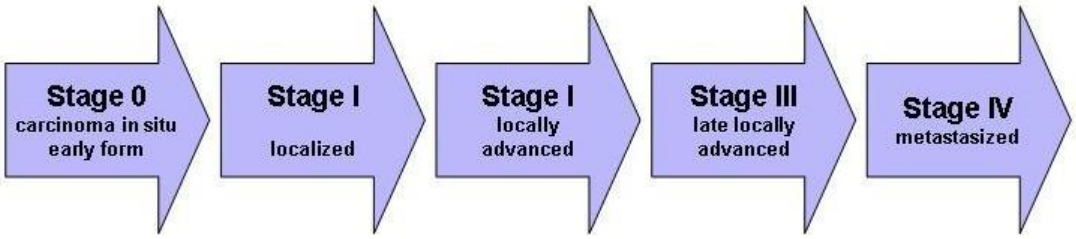


Figure 1: UICC staging of tumor progression.

CRC development

CRC arises as a result of the progressive accumulation of genetic and epigenetic changes within cells from the epithelium lining the colon and rectum [7]. Usually, CRC occurs late in lifetime. 92% of patients are diagnosed after the age of 50 [3]. More than 80% of CRC cases are regarded as sporadic whereas familial syndromes account for about 8-15% of cases [7]. There are two major forms of hereditary CRC, the familial adenomatous polyposis syndrome (FAP) and hereditary nonpolyposis colorectal cancer (HNPCC or Lynch Syndrome) [8]. FAP, a disease in which thousands of polyps develop along the colon and harbor the potential to be the founder of a tumor, was traced to a

deletion or inactivation of the adenomatous polyposis coli (*APC*) gene. Its protein (APC) plays a major role in the Wnt / β-catenin pathway (Figure 2) [9]. If APC function is lost, degradation of β-catenin is abolished. B-catenin enriches in the cytosol and translocates into the nucleus where it forms a complex with the transcription factor TCF/LEF leading to the expression of proto-oncogenes like *c-myc* or *cyclin D1*. Interestingly, 70 - 80% of sporadic CRC carry inactivating mutations of the *APC* gene, too. CRC with intact *APC* were often found to harbor activating mutations of β-catenin instead [9]. Thus, the majority of CRC cases are elicited by the aberrant activation of the Wnt / β-catenin pathways [10].

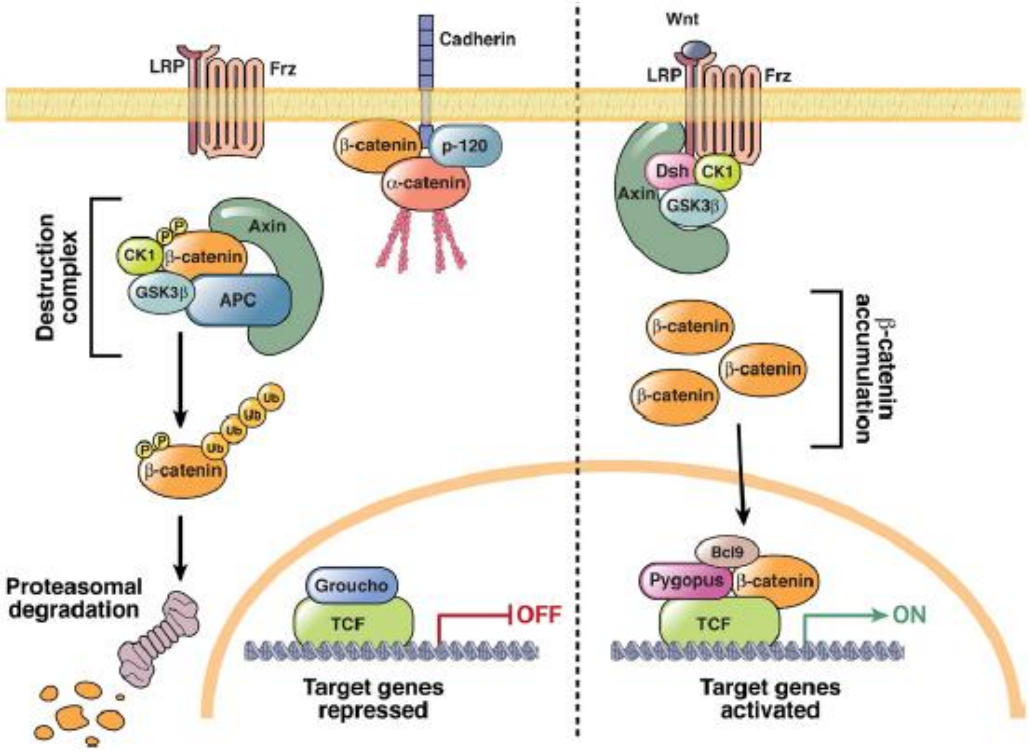


Figure 2: The Wnt signaling pathway. In the absence of a Wnt signal, the destruction complex containing APC and other proteins targets the degradation of cytoplasmic β-catenin in a proteasome-dependent manner. In the nucleus, Wnt target genes are kept silent by the repressor Groucho interacting with DNA-bound T cell factor (TCF). In the presence of a Wnt ligand, the destruction complex is inactivated through its phosphorylation dependent recruitment to the Wnt receptor Frizzled (Frz) and its co-receptor LRP. Degradation of β-catenin no longer occurs. Cytoplasmic β-catenin translocates to the nucleus, where the transcription of multiple genes is initiated through displacement of Groucho and the interaction of β-catenin with TCF/ LEF family of transcription factors. Figure taken from Pino and Chung [11].

Inactivating mutations in various mismatch-repair (MMR) genes lead to the aforementioned HNPCC disease or Lynch syndrome respectively. Silenced or inactivated MMR genes are found in about 12% of sporadic CRC as well [12]. They

maintain genomic integrity and mediate DNA damage-induced cell death. Loss of MMR function can be detected by microsatellite instability (MSI). Microsatellites (MS) are short repetitive sequences in DNA that are prone to a higher rate of mutation during DNA replication compared to other DNA regions. If MS repeat errors remain unfixed due to defects in MMR function, the length of MS can change which might lead to frameshift mutations that inactivate or alter tumor suppressor genes such as *APC*, *TP53* and *SMAD4*. MS sequences of the TGF-β receptor II (*TGFBR2*) are particularly affected by MSI. Frameshift mutations are found in the *TGFBR2* gene in 80% of CRC with MSI, leading to alterations in the TGF-β/SMAD signaling [10]. TGF-β signaling leads to the formation of a hetero-oligomeric complex of SMAD proteins that translocates to the nucleus and interacts with key transcription factors (c-jun, p300/CBP, c-myc). Moreover, several cell cycle checkpoint genes (p21, p27, p15) were found to be TGF-β signaling targets as well. Consequently, defects in TGF-β will contribute to loss of cell cycle control [13, 14].

The process of colorectal tumorigenesis is complex and involves sequential mutations of several key signaling pathways (Figure 3) [15]. The inactivation of the *APC* gene is believed to be an early event in CRC. Further mutations in oncogenes such as *KRAS* and in tumor suppressor genes like *TP53* are required for a complete malignant progression. KRAS, a member of the small G-protein family, mediates cellular signal transduction and was found to be constitutively active in up to 30-40% of CRC patients. 95% of *KRAS* mutations are found in codons 12, 13 and 61. If KRAS is constitutively active, the cellular proliferation is uncoupled from extracellular signals such as epidermal growth factors (EGFs) [16]. *TP53* encodes the transcription factor p53 that is involved in various cellular pathways including cell cycle regulation and apoptosis. It is believed that loss of p53 function might allow evasion from cell cycle arrest and apoptosis. Inactivating mutations of *TP53* often mediate the transition of adenomas to invasive cancers [17]. These mutations are located mainly in the DNA binding domain of the p53 protein and are common in about 50% of CRC cases [18].

Furthermore, the PI3K signaling cascade is frequently found to be constitutively active in CRC [19]. This is often caused by mutations in *PI3KCA*, the gene encoding the catalytic subunit of PI3K. One of the most prominent downstream targets of the PI3K pathway is Akt. Akt is critically involved in the regulation of apoptosis, gene transcription by NF-kB pathway and cell cycle progression [20]. Interestingly, immunohistochemical analyses showed that some tumors expressing high levels of

activated Akt are significantly associated with poor prognosis [21-23]. However, for CRC the strong expression of activated Akt was found to be associated with low cancer stage and favorable outcome [24].

Beyond genomic mutations increasing data imply that miRNAs are involved in the deregulation of CRC associated genes, too [25]. For instance, a decreased expression of miR-34a was found in human colon cancer tissues compared to normal tissues and might be linked to a reduced expression of TP53.

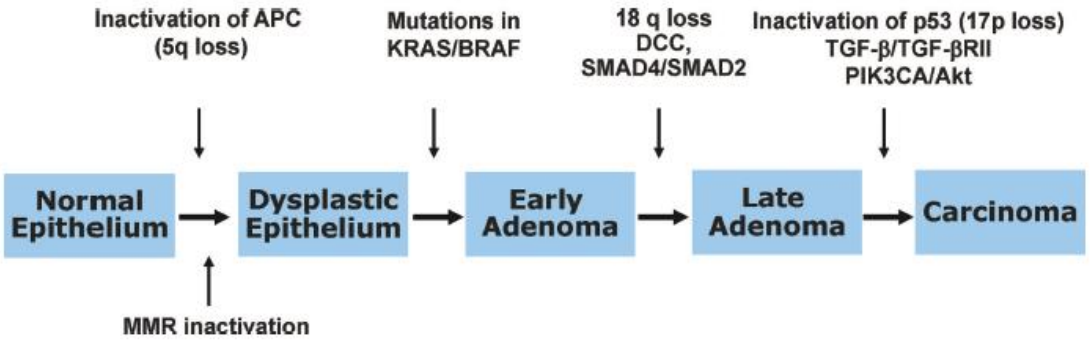


Figure 3: Multistep genetic model of CRC development. The progression of CRC is accompanied by sequential mutational events in key signaling pathways. The activation of the Wnt signaling pathway results from mutations in the *APC* gene and is an early event in tumor initiation. Further cancer progression to late adenomas and early carcinomas requires mutations in *KRAS* and *TP53* and loss of heterozygosity at chromosome 18q. Aberrant TGF-β signaling and activation of the PI3K/Akt signaling pathway through activating mutations in the *PIK3CA* gene lead to full malignancy. Figure taken from Saif and Chu. [7].

Metastasis and the epithelial mesenchymal transition (EMT)

Most common tumors, like CRC, are of epithelial origin [2]. Such carcinomas derive from epithelial cells, which are characterized by tight cell–cell interactions, the basal-apical polarization and the inability to migrate. Genetic damages lead to increased cell proliferation and to the formation of an initially benign tumor. Within a multistep process neoplastic cells accumulate further genetic and phenotypic changes and switch into malignancy. They become invasive, break through the basal lamina, enter the lymphatic vessels or penetrate the blood vasculature and reach the circulation. The mechanism of how tumor cells can disseminate and spread through the body is currently not fully understood. A widely accepted but still controversial discussed hypothesis is that tumor cells undergo an epithelial-mesenchymal transition (EMT) [26, 27, 28]. EMT is a crucial and highly conserved process in embryo development. It is

characterized by the down-regulation of cell adhesion molecules, the loss of a rigid cytoskeleton and an increased cell motility which is essential during gastrulation movements and neural crest formation. Growth factors including TGF-β, EGF, HGF/SF, bFGF and PDGF were found to be initiators of EMT. Diverse signal transduction mechanisms mediate EMT, e. g. receptor tyrosine kinases / Ras, Wnt, Notch, Hedghog and NFkB pathways [29]. It has been proposed that invasion and metastasis of carcinoma cells is linked to EMT [30]. This is supported by the finding that in most epithelial cancers E-cadherin, a cell-cell adhesion molecule, is inactivated or its expression is down-regulated. Various studies have shown that this seems to correlate with cancer grade and patient survival [31-34]. E-cadherin expression can be regulated at different levels of gene expression [35]. A major mechanism, however, is the transcriptional repression through the binding of transcriptional repressors to E-boxes located at the proximal promoter site. Most prominent transcription factors that down-regulate E-cadherin expression include Snail, Slug, SIP-1, ZEB1 and TWIST1 [36]. Various studies link these transcription factors generally to EMT [37-40]. For instance, Snail was found to down-regulate several epithelial markers such as claudins, occludins, desmoplactin and cytokeratins and up-regulate the expression of the mesenchymal markers fibronectin and vitronectin. TWIST1 has been associated with increased N-cadherin expression which was found to support cell motility.

In summary, EMT leads to large changes in the gene expression profile and might be responsible for loosing cellular assembly. Concomitantly, increased cell motility might enable tumor cells undergoing EMT to enter the vasculature and to spread over the body (Figure 4).

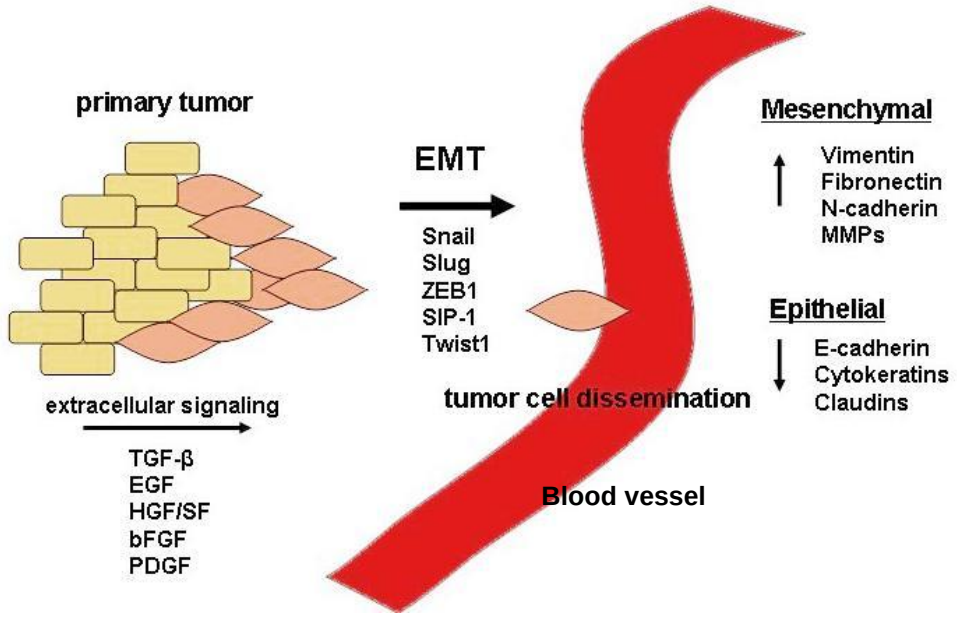


Figure 4: EMT and tumor cell dissemination. Epithelial mesenchymal transition (EMT) seems to be critically involved in tumor cell dissemination. Extracellular signaling mediates the up-regulation of EMT-inducing transcription factors in the primary tumor. The down-regulation of epithelial markers and the up-regulation of mesenchymal proteins facilitate tumor cell migration and invasion into the blood vessels.

Dissemination and colonization

Tumor cells that have detached from the primary tumor and circulate in the blood are termed CTCs. It is estimated that large numbers of CTCs enter the circulation but only a small portion is able to survive and to establish metastasis at secondary organs [41, 42]. The main portion of CTCs can not sustain the shear stress in the blood stream [43] and undergoes apoptosis. Moreover, for survival CTCs have to develop mechanisms to be protected from anoikis, which is a kind of programmed cell death that is induced when anchorage-dependent cells detach from the extracellular matrix [44]. Occasionally observed CTC clusters or multicellular tumor cell aggregates [45] might be a way to protect CTCs from mechanical burdens and anoikis. Furthermore, CTCs need to be protected from immunosurveillance. Besides a decreased expression of MHC I proteins [46] it has been reported that CTCs might avoid NK cell-mediated lysis through the aggregation with platelets [47, 48]. In addition, Pawelek et al. have suggested that CTCs might fuse with bone marrow-derived cells (BMDC) in the blood [49]. The BMDC – CTC fusion might increase CTC survival and could provide an explanation for acquired mesenchymal traits in CTCs.

Liver metastases are pre-dominantly observed in CRC [50-52]. CTCs, which are shed from the primary tumor, reach the liver through the mesenteric and portal venous system. The majority of the CTCs is filtered by the liver before they are able to enter the systemic circulation. According to a cascade theory [50] of cancer progression that suggests the step-wise progression of malignancy, the CTCs that are trapped in the liver grow out to metastases and increase the amount of CTCs in the blood. This might explain why CTCs in peripheral blood are primarily found in patients with overt metastases [53]. In contrast to that hypothesis it is also conceivable that numerous CTCs pass the liver and distribute throughout the body. In this alternative scenario the usually observed homing of CTCs to liver or lung would be attributed to adhesion molecules and certain cytokines. Already in 1889 Stephen Paget hypothesized that metastasis is not a random process and presumed that there must be a strong impact of the interaction between tumor cells and the microenvironment [54]. Paget developed the "seed and soil" hypothesis claiming that certain tumor cells (the seed) metastasize to certain organs (the soil). It is widely accepted that tumors and CTCs are a very heterogeneous cell population. The question which properties of seed and soil efficiently support metastasis is still largely unknown. However, the local expression of chemoattractants and cell adhesion molecules might explain the organ tropism of disseminated cancer cells. It was found that the expression of the chemokine receptor CXCR4 on breast cancer cells guides them to tissues with high expression levels of the CXCR4 ligand SDF-1. These tissues include bone marrow, liver and lung [55]. In addition, a receptive microenvironment seems to be a prerequisite to successfully establish a metastatic lesion at a distant site. It was found that VEGFR-1^+ bone marrow-derived hematopoietic progenitor cells can be mobilized to the invasive front of the primary tumor as well as to the designated target organs of metastasis where they support angiogenesis and tumor progression [56]. Already before the arrival of tumor cells VEGFR-1^+ bone marrow-derived hematopoietic progenitor cells together with endothelial and stromal cells prepare the pre-metastatic niche as the future site of metastasis through modification of the extracellular matrix and complex chemokine signaling. Once, CTCs/DTCs reach the pre-metastatic site, they extravasate from the blood vessels and become sessile again. It is supposed that the process of EMT is reverted and the epithelial traits need to be re-gained to from metastases [57]. Probably, the DTCs stay initially in a dormant state. Metastatic relapse in patients with CRC can occur years after diagnosis and resection of the primary tumor [58]. Therefore, it is

assumed that many cancer patients harbor dormant tumor cells that even resist chemotherapy due to their low proliferation rate, appropriate detoxifying enzymes and presumable stem cell properties. Currently unknown factors can trigger tumor cell proliferation again. In this case the metastatic niche undergoes an angiogenic switch by recruiting endothelial progenitor cells. Finally, increasing vascularization enables the progressive growth to marcometastases.

Tumor stem cells

Although a tumor develops from a single, mutated cell, tumor cells within a tumor are not identical [59]. One explanation could be the genomic instability of tumor cells [60] and a constant selective pressure for cells to adapt to the tumor microenvironment. Alternatively, it was proposed that a hierarchy exists in tumors with only a minority of cells that are capable of regenerating the tumor. These cells are termed cancer stem cells [61]. The bulk of the tumor, however, is formed by cells with a limited capacity to divide. Currently it is not clear if tumor stem cells originate from normal stem cells or from progenitor cells with a blocked differentiation program or even from differentiated cells that somehow acquired self renewing capacity. The hypothesis that cancers like all other tissues originate from stem cells is not new and is supported by the finding that only a small subpopulation of cancer cells is able to induce tumor growth in mouse models. Biomarkers to identify and define cancer stem cells, among them CD44, CD133 and CD26, are still not standardized and under intensive investigation [62, 63]. Nevertheless, the cancer stem cell hypothesis is still controversial [64-66]. Critics state that it might be an *in vitro* assay-associated artifact. Supporting evidences rely mainly on xenotransplantion studies in immunosuppressed mice. Furthermore, they allude that the low numbers of human tumor cells producing tumors in mice arise from difficulties of the tumor cells to adapt to the novel and foreign microenvironment. Hence, common assays rather select for cells that are tumorigenic in mice but neglect that tumor formation depends on complex interactions with extracellular matrix components and multiple non-tumor cells that might not be present in the xenotropic mouse model.

Is metastasis initiated by tumor stem cells?

If tumor stem cells exist, then they might be already present in benign lesions. Consequently, additional features have to be acquired to develop malignant growth. An intriguing concept of malignant tumor progression was presented by Brabletz and colleagues[67]. They propose the existence of a mobile cancer stem cell fraction in tumors and metastases they term migrating cancer stem cells (MCS). According to their hypothesis, these cells might originate from stationary cancer stem cells that additionally gained cellular mobility through transient acquisition of EMT. The concept comprises that a migrating front of MCS extends the tumor mass at the tumor-host interface and the formation of metastases is caused by long distance migration of MCS.

Cancer progression models

Early or late dissemination of tumor cells

The resection of the primary tumor is often not curative in CRC since tumor cells might have already disseminated and are spread throughout the body [68]. Systemic treatments to prevent the formation of distant metastases were developed to target DTCs [69]. However, DTCs are poorly characterized and we know hardly anything about the metastatic founder cells so far. From this it follows that predictions about the properties of DTCs require at least models of cancer progression. In particular it is important to elucidate if primary tumors can be used to predict the success of systemic treatments to defeat DTCs (Figure 5). Two basic models of metastasis are discussed [70]. Basically, they are focused on the question if tumor cell dissemination is an early or late event in tumor progression. The linear progression model states that cancer cells disseminate not until the tumor reaches full malignancy. Consequently, DTCs resemble the primary lesion. The second model suggests a parallel progression of DTCs and the primary tumor. This means that DTCs leave the primary lesion already at an early point of time. In this scenario DTCs develop independently and divergent from the primary tumor. There are arguments for both models. The linear progression model is supported by the finding that the appearance of metastases correlates with tumor size and that an early resection of the primary lesion prevents dissemination of cancer cells and increases the patients' survival. However, the model can not explain the occurrence of metastases in early stage cancer or of metastases of unknown primary site. Moreover, based on the assumption that primary tumors and metastases have a comparable growth rate, one can

estimate the point of time when a metastasis started to grow. Data from registries, that report the time from resection to the appearance of distant metastasis, contradict the late dissemination of tumor cells because metastases would be expected to emerge much later in time. In addition, genetic evidence favors the parallel progression model. Techniques for whole genome analysis of single DTCs revealed that significant fewer genetic aberrations can be found in DTCs compared to primary tumor cells indicating an early dissemination of tumor cells. Furthermore, genomic disparities in matched primary tumors and manifested metastases rather support the parallel progression model.

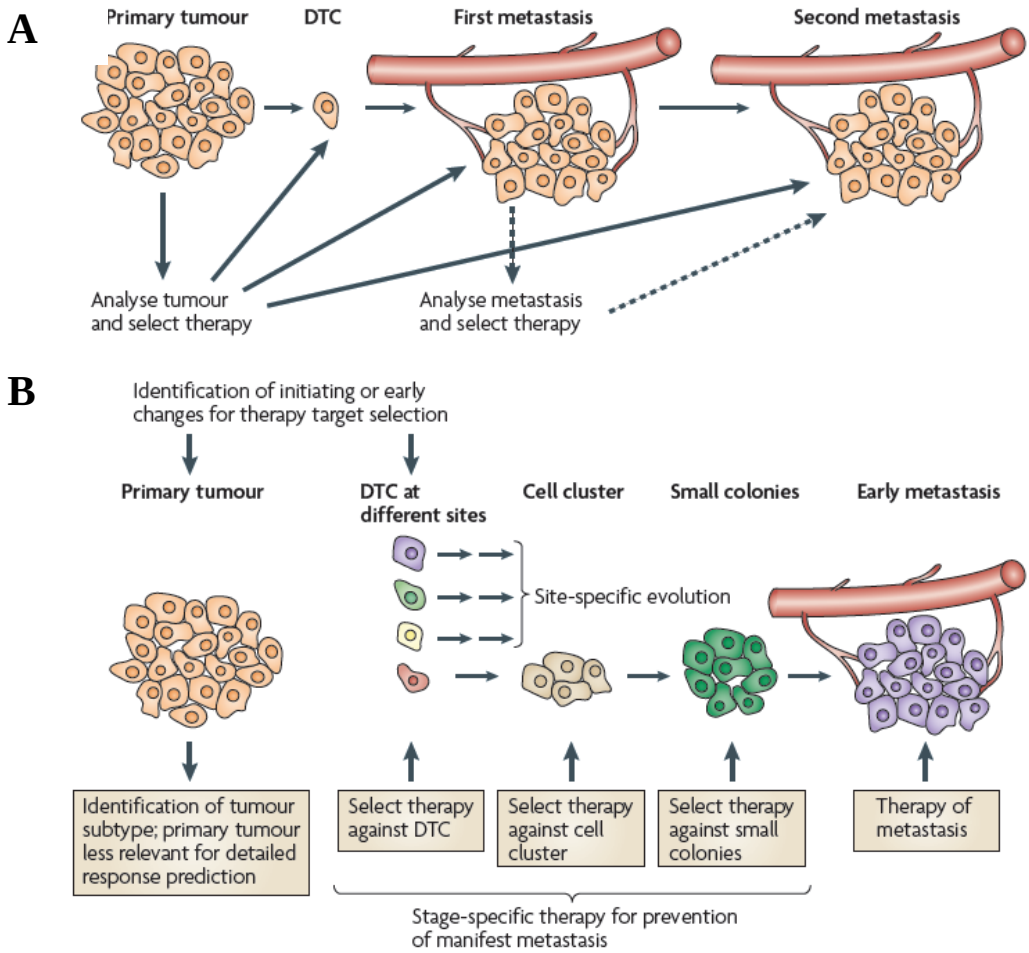


Figure 5: Early or late dissemination of cancer cells. **A** The late dissemination and metastatic cascade model predicts that disseminating tumor cells (DTCs) resemble the most aggressive clone and therefore the tumor at time of diagnosis. Therapies can be selected on the basis of analysis of the primary tumor. **B** Parallel progression complicates therapy selection. Genomes may vary between different sites and undergo independent progression. Specific genetic changes might be selected at different sites and different targeting strategies have to be applied for DTC, small clusters and small colonies. Figure and description was taken from Klein [70].

Disseminated tumor cells

Definition

The term minimal residual disease (MRD) comprises single tumor cells and micrometastases throughout the body. According to the location in the body further distinctions are required:

- Disseminated tumor cells (DTCs) comprise tumor cells that are found in lymph nodes or in the bone marrow.
- Circulating tumor cells (CTCs) are located in the blood.

Methods of detection

Metastatic spread due to tumor cell dissemination is the most threatening aspect of cancer. To increase our knowledge about the metastatic process several techniques have been developed to capture CTCs in blood and DTCs in bone marrow and lymph nodes for enumeration and characterization. The presence of lymph node - associated DTCs is usually confirmed immunohistologically by epithelial specific staining, wheras diverse techniques for the detection of CTCs and bone marrow-associated DTCs exist. Techniques for CTC and DTC detection are highly analog; therefore, this chapter is just focused on CTC identification methods.

Since CTCs are very rare and hard to find in a strong background of leukocytes (often less than one in 10^6 peripheral blood mononucleated cells (PBMC)), sophisticated approaches are needed for CTC detection [71, 72]. Most protocols apply a density gradient centrifugation in combination with an immunomagnetic isolation step to enrich CTCs from a blood sample. However, CTCs seem to be very heterogeneous in their morphology and in the expression of cell-surface antigens which probably leads to a bias in all used enrichment methods. Most data were obtained from the CellSearch™ (Veridex) system [73] which represents the only FDA- approved device for CTC detection in colorectal, prostate and breast cancers that allows the enumeration of CTCs in whole blood specimens of cancer patients. CTCs can be distinguished from leukocytes due to the expression of epithelial protein markers which are normally absent on blood cells. The purification / detection process starts with the immunomagnetical enrichment of EpCAM-expressing CTCs with the aid of an EpCAM-conjugated ferrofluid. Afterwards the specimens are stained with a nuclear stain to prove cellular integrity and fluorescent antibody conjugates directed against CD45, cytokeratin 8, 18 and 19. A trained operator can distinguish cytokeratin-positive

CTCs from residual CD45- positive leukocytes to determine the CTC amount in a 7.5 ml blood sample (Figure 6).

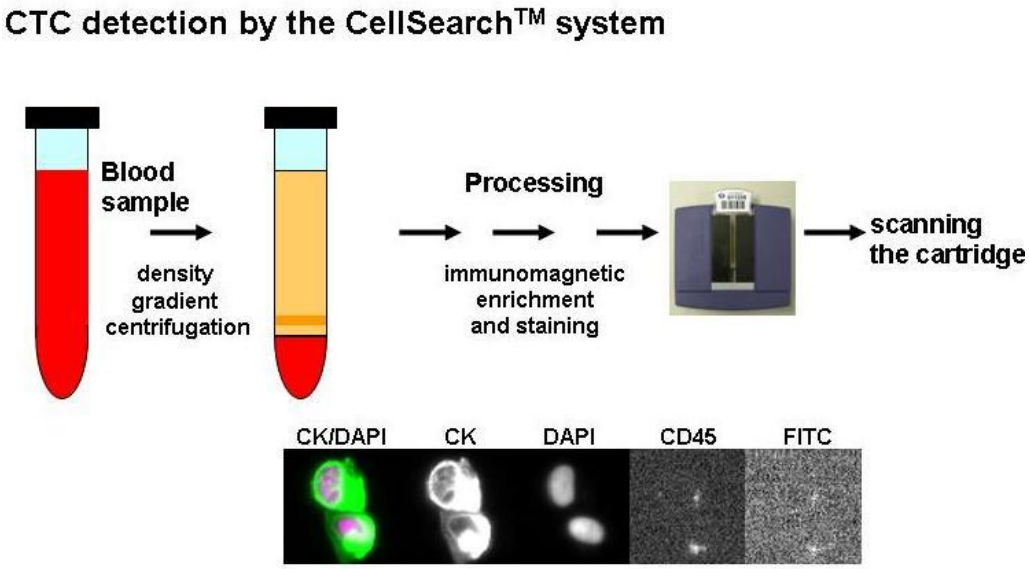


Figure 6: The CellSearch. The CellSearch™ system is used for the detection of CTCs. The figure summarizes the processing of a blood sample as explained in the text.

Several other techniques for CTC detection have been developed recently. Most of them rely on EpCAM and cytokeratin expression of the target cells. For instance, the CTC chip is a microfluidic device that captures EpCAM-expressing CTCs within an array of antibody coated microposts [74]. Subsequent cytokeratin staining enables the identification of CTCs. EpCAM expression of target cells is also required for the automated cell separator MagSweeper™ [75] that is designed to perform efficient immunomagnetic CTC enrichment. Other surface-marker-based CTC detection devices include standard flow cytometry and laser scanning cytometric based systems, such as Maintrac™, Ikonoskope™ and Ariol™.

A different approach to isolate CTCs is the ISET system [76]. Since CTCs usually are larger than peripheral blood leukocytes, it is possible to separate them with the ISET system by size filtration. However, the method is controversial because the CTC size is highly variable and sometimes comparable to leukocytes [77].

Another technique, the EPISPOT (Epithelial immunoSPOT), was designed to detect viable CTCs [78]. Leukocyte-depleted blood or bone marrow specimens, that potentially contain CTCs, are cultured to accumulate a sufficient amount of marker proteins. By adapting the enzyme linked immunospot technology CTC detection is getting feasible.

Other techniques use polymerase chain reaction (PCR) based systems. The Adnatest™ method relies on the immunomagnetic CTC isolation and the subsequent detection of

tumor marker expression in a multiplex PCR [79]. A big drawback of that method is that CTCs can not be quantified. The lack of quantification is caused by the impossibility to determine from how many cells the detected transcripts are derived.

In summary, despite several CTC detection methods are available a standardized CTC detection system without bias is still lacking. CTCs conserve characteristics of the epithelial cancer tissue from which they originate. This is the basis for the common detection methods; however, it is also the main point of critics. In the course of EMT and the dissemination process CTCs might down-regulate the expression of epithelial markers, such as EpCAM, a protein that is also involved in cell adhesion. Standardized and more unbiased methods for CTC detection would be favorable to get a comprehensive insight in tumor cells shedding.

Clinical significance of CTCs and DTCs in CRC

Routine histopathological methods have confirmed that lymph node involvement is one of the most important prognostic factors for colon cancer patients [80]. However, these methods frequently miss smaller lesions and micrometastases. Increasing sensitivity of detection due to molecular detection methods revealed that lymphatic spread of tumor cells seems to be common already in early stage CRC. Since lymph node - associated CTCs are found in N_0-staged lymph nodes (analyzed as negative in conventional pathology) in up to 86% of cases, it is not clear to what extent micrometastases and single lymph node - associated CTCs contribute to a worse prognosis in CRC [81 82 83].

Numerous studies addressed the question if the detection of CTCs and DTCs are a prognostic factor for progression free survival and overall survival in CRC [84]. Due to an increasing amount of data obtained from the CellSearch™ system the case for metastatic CRC (mCRC) seems to be quite clear [85 86, 87]. Cohen showed that CTC counts above the baseline of 2 CTCs per 7.5 ml of peripheral blood indicate an about 50% reduced progression free survival and overall survival in mCRC patients (Figure 7).

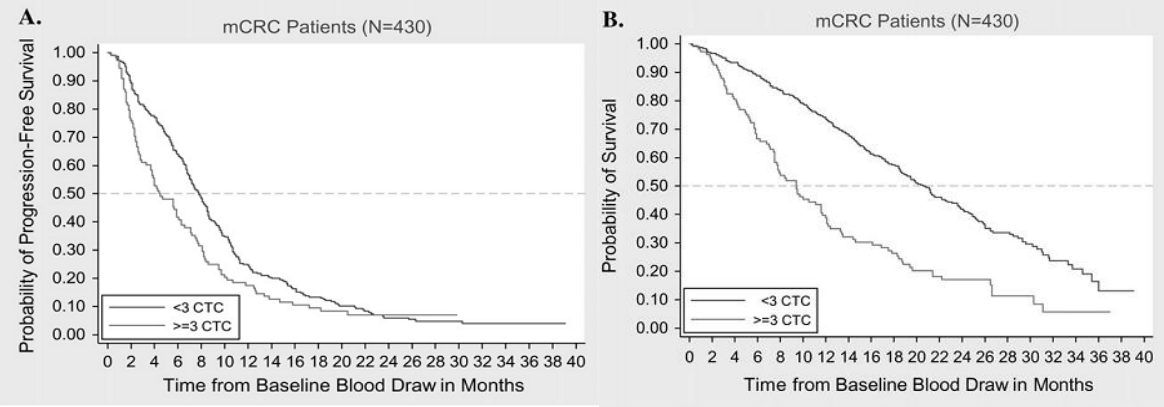


Figure 7: CTC count and prognosis in mCRC. Progression-free survival (panel A) and overall survival (panel B) for metastatic CRC patients by favorable (<3) versus unfavorable (≥3) baseline CTC count. Figure was taken from Cohen et al. [86]

However, the impact of CTCs and DTCs in non-metastasized CRC patients is still controversial. The main problem is that the available studies vary dramatically in study design and in the applied methods. Standardized guidelines are required to obtain comparable study results.

Only a few studies address the prognostic significance of DTCs in the bone marrow of CRC patients [84]. It is supposed that the bone marrow might serve as a reservoir of DTCs and would be a good indicator for MRD. Lifespan of CTCs in the blood is short [88] and the fact that CTCs can still be found in the blood months after removal of the primary tumor, indicates that CTCs might circulate among different metastatic sites [89]. The bone marrow might be a source of CTCs even years after curative resection. DTCs usually stay there in dormancy until they are released again due to unknown stimuli. Since the available studies are very heterogeneous in sample size, staining methods and follow-up-times, further studies are required to finally assess the role of DTCs in CRC [90].

Molecular characterization of DTCs and CTCs

Understanding the process of metastasis requires the molecular analysis of CTCs and DTCs. However, their rareness is the major difficulty for a detailed characterization. Fortunately, along with improvements in single cell techniques the characterization of the metastases precursor cells is becoming more and more feasible.

Immunocytochemistry

The phenotypic heterogeneity of DTCs isolated from bone marrow samples was shown by immunocytochemical double or triple stainings [91] (Figure 8). DTCs detected in the bone marrow are often dormant which was indicated by the low expression of proliferation markers such as Ki67, p120 and PCNA. This might explain the appearance of clinical metastases late after the curative resection of the primary tumor [58]. With regard to immune escape mechanisms the down-regulation of major histocompatibility complexe I (MHC I) was found on DTCs of breast cancer patients [92]. Recent studies also report that the over-expression of Her2, a receptor for EGFR, in breast cancer-associated DTCs and CTCs is linked to a poor clinical outcome [93]. This finding might support the use of trastuzumab^{TM}, a humanized anti-Her2 monoclonal antibody, for systemic treatment in Her2 over-expressing breast cancer patients. Furthermore, urokinase-type plasminogen activator receptor (uPar) expression on DTCs correlates with metastatic relapse in gastric cancer [94]. It was speculated that signaling mediated by Her2 and uPar might be involved in growth induction in dormant DTCs. Since heterogeneity of CTCs and DTCs becomes more and more evident [45], the search for the metastasis founder cell fraction is proceeding. Based on CD44, CD24 and ALDH1 the existence of a potential cancer stem cell subpopulation was reported for breast cancer-associated CTCs [95]. However, so far nothing is known about the prognostic significance of this finding.

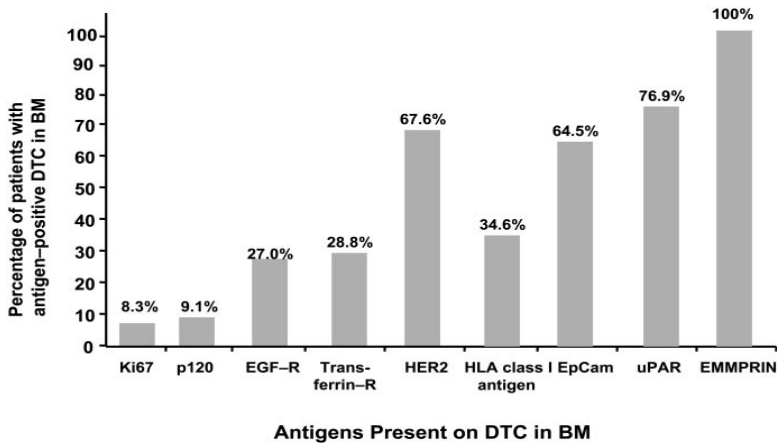


Figure 8: Heterogeneity of DTCs. Phenotypic profile of DTCs in bone marrow determined by immuno-cytochemical double staining using anti-cytokeratin antibodies for tumor cell detection. Figure was taken from Pantel et al. [91]

Genomic characterization

The genomic analyses of DTCs and CTCs were initially necessary to confirm the malignant origin of cytokeratin-positive stained cells in the bone marrow. By means of fluorescence-in-situ-hybridization [96] it was possible to show chromosomal aberrations in CTCs / DTCs. When later whole genome amplification techniques became available, single cell comparative genomic hybridization (CGH) was applied to screen for chromosomal gains and losses. It was reported that DTCs generally display fewer genetic aberrations than their matched primary tumors in breast, prostate and esophageal cancer supporting the theory of an early dissemination of tumor cells [97]. Growing evidence indicates that DTCs might develop independently from the primary tumor. This may lead to primary tumors and metastases with different mutations in oncogenes or tumor suppressor genes. For instance, disparate *KRAS* mutations of tumors and matched metastases were observed in up to 60 % of CRC [98]. Since constitutively activating *KRAS* mutations in CRC have been associated with failure of anti-EGFR therapy, the molecular profile of CTCs/ DTCs is becoming increasingly important to defeat systemic cancer and to predict therapy response [99].

Transcriptome analysis

Array-based gene expression analysis could provide crucial information about the dissemination and survival mechanisms of CTCs and DTCs and might also permit the identification of novel therapeutic targets fighting systemic cancer. However, these approaches are mainly limited by the low number of CTCs / DTCs and impurity of the applied isolation strategies. Most of the available studies were solely able to perform RNA expression analyses of CTC or DTC - enriched cell fractions in a background of several thousands of leukocytes [100-102]. Despite these limitations Watson et al. provided supporting evidence for EMT in DTCs by means of global gene expression profiling [102]. They detected the expression of TWIST1, a putative inducer of EMT, in DTCs in early stage breast cancer patients and found TWIST1 expression correlated with early disease relapse.

Klein et al. [103] reported the transcriptome analysis of single micrometastic cells from bone marrow samples. With the help of a sophisticated RNA extraction and subsequent cDNA amplification strategy it was possible to compare the gene expression profiles of DTCs and to conclude that most of DTCs adapt a dormant cell cycle status. Additionally, they detected the expression of EMMPRIN (extracellular matrix

metalloproteinase inducer) in the analyzed micrometastatic cells. EMMPRIN is thought to facilitate tumor invasion by stimulation of matrix metalloproteinases. Immunocytostaining of cytokeratin positive DTCs of diverse cancers revealed EMMPRIN expression in 82 % of DTCs indicating their invasive properties.

Results

The objective of this thesis was to characterize CRC-associated CTCs on the transcriptional and genomic level. Following the trail of hematogenous tumor cell dissemination we recorded the CTC amount and distribution in patients' blood obtained from local tumor draining veins and central veins. CTCs that have been detected were isolated for their genomic characterization with regard to point mutations, MSI and chromosomal instability. In addition, intact and freshly processed CTC samples were required to enable mRNA expression profiling. Therefore, we had to establish a reliable CTC enrichment and isolation protocol. Finally, a mouse model of metastatic CRC was developed. The CTCs that were obtained from the mouse model ought to provide further insights into the dynamic regulation of mRNA expression involved in tumor cell dissemination.

Compartmental differences of CTCs in CRC

Due to the clinical significance and prognostic value of CTCs in CRC, we were deeply interested if the presence and the detection rate of CTCs differs among different blood compartments. For that purpose we used the US Food and Drug Agency (FDA) approved CellSearch™ system to quantify the amount of CTCs in our patient cohort and to investigate the distribution of CTCs in the mesenteric venous blood compartment (MVBC) compared to the central venous blood compartment (CVBC).

CRC - associated CTCs that are shed from the primary tumor reach the liver via the mesenteric and portal venous system before they enter the systemic circulation [50]. Our study ought to provide data about primary hematogenous CTC dissemination and systemic hematogenous spread with regard to an assumed CTC filtering function of the liver. Furthermore, the presence of CTCs was correlated to various clinicopathological parameters. The study was conducted and supervised by our group member Dr. Rahbari and has been published [104].

Blood samples of 200 CRC patients who underwent surgical resection for CRC in the department of surgery of the University Clinic Heidelberg between May 2009 and April 2011 were enrolled prospectively. Table 1 gives an overview about the study design and clinicopathologic characteristics of the study population.

Central venous blood of 200 patients was analyzed for the presence of CTCs whereas 83 % of the analyzed blood samples showed no detectable CTCs. Univariate analysis (Table 2) confirmed that the number of patients with CTCs as well as the CTC amount correlated with the stage of disease. CTCs were detected in 10.2 % (19 of 142) of patients with UICC stage I-III and in 35.8 % (19 of 53) of UICC stage IV patients ($p<0.0001$). Stage I-III patients showed a mean amount of 0.2 (range: 0-5) CTCs per 7.5 ml blood sample and 2.4 (range: 0-83) CTCs were detected in stage IV specimens ($p<0.0001$). Serum levels of CEA and CA 19-9 were associated significantly with detection of CTCs in central blood (Table 2).

Regarding the detection of CTCs in MVBC data of 80 patients (40%) were available. A significant higher rate of patients with CTCs in the MVBC (35%, n=28) compared to the CVBC (17.5%, n=14) ($p = 0.01$) was found. Furthermore, the average number of CTCs was higher in the MVBC (mean 1.5, median 0, range: 0-32) than in the CVBC (mean 0.3, median 0, range 0-5) ($p = 0.006$) (Figure 9). CTCs in the CVBC were more frequently found ($p = 0.01$) and with higher number of tumor cells ($p = 0.01$) in patients with detectable CTCs in the MVBC. However, the presence of CTCs in the MVBC did not correlate with the staging of the disease (Table 3). Univariate association of clinicopathological variables with CTC presence and amount in MVBC showed that CTCs in the MVBC were detected at a higher rate ($p = 0.02$) and with higher quantity ($p = 0.01$) in patients with colon compared to patients with rectal cancer (Table 2). Correlations of tumor cell detection in MVBC to other assessed parameters were not found.

In summary, we found that the number of patients with CTCs and the amount of CTCs was significantly correlated with the stage of disease. In addition, a significant higher rate of patients with CTCs in tumor-draining venous blood compared to the central venous blood was detected.

Tabel 1: Clinicopathological characteristics of the study population. [104]

Clinicopathological parameter	N (%) or median (range)
Age	65 (27-86)
Sex	
Male	129 (64.5)
Female	71 (35.5)
Site of disease	
Colon	96 (48)
Rectum	104 (52)
T stage	
T0	5 (2.5)
T1	13 (6.5)
T2	50 (25.0)
T3	107 (53.5)
T4	25 (12.5)
N stage	
N0	114 (57.0)
N1	52 (26.0)
N2	34 (17.0)
UICC stage	
0	5 (2.5)
I	45 (22.5)
II	47 (23.5)
III	50 (25)
IV	53 (26.5)
Tumor differentiation	
Moderate (G2)	112 (77.2)
Poor (G3)	33 (22.8)
CEA	
Elevated(> 2.5 U/l)	65 (34.4)
CA-19-9	
Elevated (> 37 U/l)	29 (15.3)

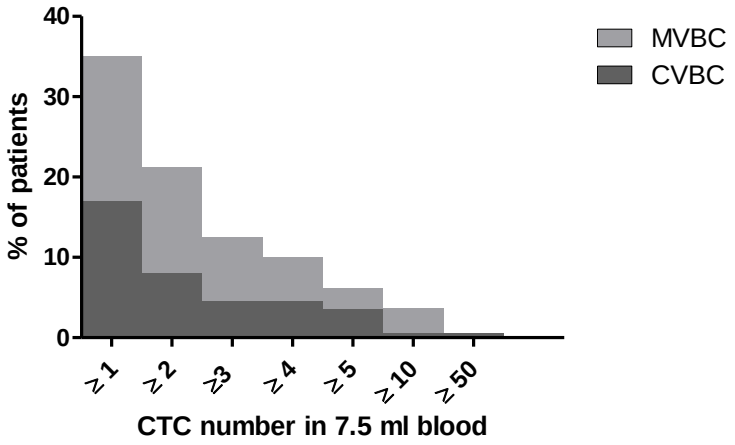


Figure 9: CTC amount in central venous blood compartment (CVBC) and mesenteric venous blood compartment (MVBC) of CRC patients. 83% of patients were free of detectable CTCs in CVBC and 65% of patients were without CTCs in MVBC.

Table 2: Clinicopathologic correlations. Associations of clinicopathologic variables with tumor cell detection in the central and mesenteric venous blood compartment of patients with CRC (stage I – IV) [104]

	Central venous blood				Mesenteric venous blood			
Characteristics	**Pt. with CTC**		**No. of CTC**		**Pt. with CTC**		**No. of CTC**	
Age		0.13		0.15		0.54		0.57
≤ 65	20 (21.3)		1.2 (0; 0 – 83)		16 (38.1)		1.4 (0; 0 – 14)	
> 65	14 (13.3)		0.3 (0; 0 – 5)		12 (31.6)		1.5 (0; 0 – 32)	
Gender		0.25		0.23		0.09		0.04
Male	19 (14.7)		0.3 (0; 0 – 5)		14 (28.0)		1.0 (0; 0 – 14)	
Female	15 (21.1)		1.6 (0; 0 – 83)		14 (46.7)		2.1 (0; 0 – 32)	
Site of disease		0.31		0.23		0.02		0.01
Colon	19 (19.8)		1.4 (0; 0 – 83)		17 (48.6)		2.4 (0.5; 0 – 32)	
Rectum	15 (14.4)		0.2 (0; 0 – 4)		11 (24.4)		0.7 (0; 0 – 14)	
Stage of disease		<0.0001		<0.0001		0.04		0.09
I – III	15 (10.2)		0.2 (0; 0 – 5)		16 (28.1)		1.4 (0; 0 – 32)	
IV	19 (35.8)		2.4 (0; 0 – 83)		12 (52.2)		1.6 (1; 0 – 14)	
T stage		0.07		0.04		0.65		0.48
T0-2	7 (10.3)		0.1 (0; 0 – 1)		18 (33.3)		1.0 (0; 0 – 14)	
T3-4	27 (20.4)		1.1 (0; 0 – 83)		10 (38.4)		2.4 (0; 0 – 32)	
N stage		0.12		0.07		0.52		0.45
N0	14 (13.2)		0.2 (0; 0 – 2)		13 (31.7)		1.3 (0; 0 – 32)	
N1/2	20 (21.3)		1.4 (0; 0 – 83)		15 (38.5)		1.6 (0; 0 – 14)	

Tumor differentiation		0.33		0.26		0.13		0.06
Moderate (G2)	16 (14.3)		0.3 (0; 0 – 5)		14 (30.4)		0.8 (0; 0 – 14)	
Poor (G3)	7 (21.2)		0.7 (0; 0 – 5)		6 (54.5)		5.0 (1; 0 – 32)	
Resection margin		0.004		0.003		0.09		0.19
R0	28 (14.9)		0.7 (0; 0 – 83)		24 (32.4)		1.4 (0; 0 – 32)	
R1/2	6 (46.1)		1.0 (1; 0 – 5)		4 (66.7)		1.8 (1; 0 – 8)	
Neoadjuvant therapy		0.81		0.89		0.62		0.45
Yes	10 (16.7)		0.2 (0; 0 – 4)		7 (25.9)		0.5 (0; 0 – 5)	
No	18 (15.2)		0.34 (0; 0 – 5)		12 (31.6)		1.8 (0; 0 – 32)	
CEA		0.0001		<0.0001		0.16		0.23
Normal	11 (8.9)		0.1 (0; 0 – 4)		16 (31.4)		1.6 (0; 0 – 32)	
Elevated	20 (30.7)		2.0 (0; 0 – 83)		12 (48.0)		1.3 (0.5; 0 – 14)	
CA 19-9		<0.0001		<0.0001		0.10		0.15
Normal	18 (11.3)		0.2 (0; 0 – 5)		22 (33.3)		1.5 (0; 0 – 32)	
Elevated	13 (44.8)		3.9 (0; 0 – 83)		6 (60.0)		1.2 (1; 0 – 4)	

Table 3: Detection of CTCs in mesenteric and central venous blood compartments and correlation to clinical staging. [104]

	Pt. with CTC	*P*	No. of CTC	*P*
Central venous blood vs. mesenteric venous blood (n = 80)		0.01		0.006
Central venous blood	14 (17.5)		0.3 (0; 0 – 5)	
Mesenteric venous blood	28 (35.0)		1.5 (0; 0 – 32)	
Detection of CTC in mesenteric venous blood (n = 80)		0.01		0.02
Patients with CTC in central venous blood	9 (64.3)		2.0 (1; 0 – 14)	
Patients without CTC in central venous blood	19 (29.2)		1.3 (0; 0 – 32)	
Detection of CTC in central venous blood (n = 80)		0.01		0.01
Patients with CTC in mesenteric venous blood	9 (32.1)		0.5 (0; 0 – 5)	
Patients without CTC in mesenteric venous blood	5 (9.8)		0.1 (0; 0 – 2)	
Stage of disease (UICC)				
Central venous blood (n = 200)		0.0007		0.23
0	1 (20.0)		0.2 (0; 0 – 1)	
I	3 (6.8)		0.1 (0; 0 – 1)	
II	6 (12.7)		0.2 (0; 0 – 2)	
III	5 (10.0)		0.3 (0; 0 – 5)	
IV	19 (35.8)		2.4 (0; 0 – 83)	
		0.27		0.68
Mesenteric venous blood (n = 80)				
0	1 (33.3)		0.3 (0; 0 – 1)	
I	6 (37.5)		2.7 (0; - 32)	
II	5 (23.8)		0.7 (0; 0 – 4)	
III	4 (23.5)		1.3 (0; 0 – 14)	
IV	12 (52.1)		1.6 (1; 0 – 14)	

Establishment of a reliable CTC enrichment and detection method

We aimed to establish a protocol for an efficient enrichment and detection of CTCs that should provide the opportunity to analyze gene expression of the captured CTCs. The previously described study on CTC enumeration with the CellSearch^{TM} system showed that CTCs are hardly to find in blood samples of CRC patients. Therefore, high efficiency and sensitivity of the enrichment method is inevitably necessary.

Magnetic beads fail to directly isolate EpCAM - expressing tumor cells

A first attempt was designed to directly target EpCAM-expressing CTCs in the blood with the help of magnetic beads. Hence, DynaBeads coated with a set of different antibodies directed against EpCAM (Moc31, VU1D9, HEA125, BerEP4, KS1/4) and CEA (B1.1/CD66) were used. All antibodies were titrated to obtain the optimal concentration and ratio. Through the use of different antibody clones that were expected to target different epitops we supposed to increase the efficiency of the CTC enrichment. Spiking experiments with HT29 CRC cell line showed that direct enrichment with anti-EpCAM-coated magnetic beads is more efficient after density gradient centrifugation (Ficoll) compared to a direct use of magnetic beads in whole blood specimen. Efficiency was determined by RT-PCR for EpCAM and CK20 expression (data not shown). However, the direct enrichment of spiked HT29 from the peripheral blood mononuclear cells (PBMC) was not superior to direct RNA extraction of the spiked PBMC pellet. Moreover, the high detection limit of 1000 spiked cancer cells and the strong carry-over of leukocytes made the method too inefficient for further considerations (data not shown).

Leukocyte depletion with magnetic beads results in 10fold enrichment

CD45 is a pan-leukocyte marker, which is normally absent on epithelial cells. It was decided to use CD45 to deplete leukocytes in blood samples and to enrich the residual CTC fraction. Figure 10 shows that the differential CD45 and EpCAM expression of leukocytes and tumor cells enables efficient leukocyte depletion and concomitant tumor cell enrichment. Even when low numbers of tumor cells were spiked an enrichment of more than 10 fold was possible (Figure 10). However, leukocyte carryover is still

strong; therefore, further techniques are required to obtain pure CTC samples for mRNA expression analysis.

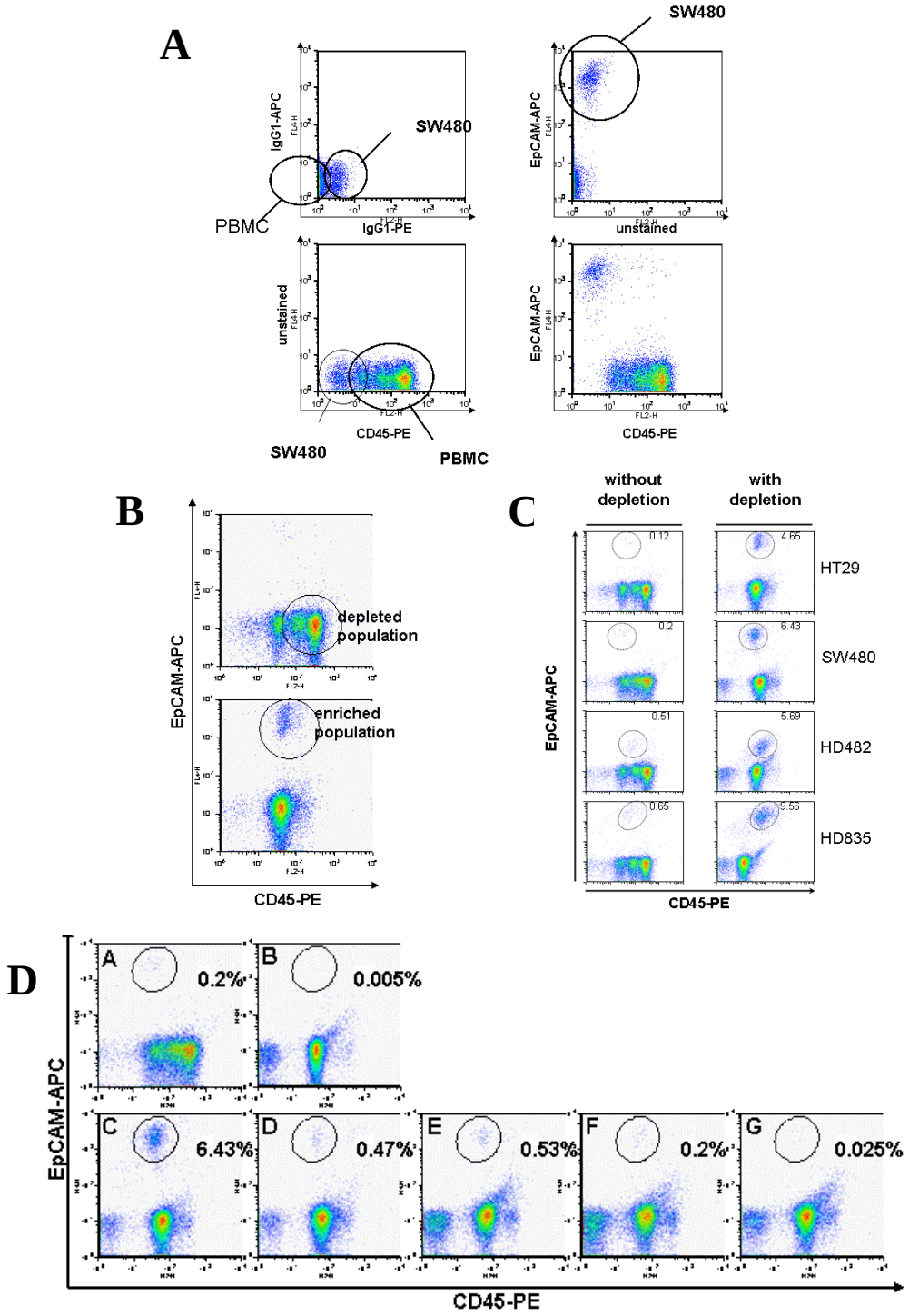


Figure 10: Surface antigen expression allows identification of CTCs A: PBMC were spiked with SW480 colorectal tumor cells and stained for EpCAM and CD45. Tumor cells and leukocytes are detectable in clearly separated populations. **B:** Leukocytes that highly express CD45 are preferentially depleted by anti-CD45 coated magnetic beads leading to tumor cell enrichment. **C:** Different CRC cell lines (HT29, SW480, HD482, and HD835) were spiked in PBMC and subsequently enriched to demonstrate the efficiency of the described method. **D:** Low amounts of tumor cells were spiked in PBMC. The tumor cell enrichment by CD45 based leukocyte depletion was tested. **DA:** undepleted control, **DB:** unspiked, depleted control **DC:** 0.5 % spiked, depleted, **DD:** 0.05% spiked, depleted, **DE:** 0.01% spiked, **DF:** 0.005% spiked, **DG:** 0,001% spiked

Fluorescence-activated cell sorting (FACS) fails to isolate CTCs

FACS sorting was considered to isolate CTCs out of leukocyte - depleted and EpCAM / CD45 stained blood samples. Therefore, leukocyte specimens were spiked with 0.5% tumor cells. After leukocyte depletion with the help of magnetic beads and fluorescence staining a defined number of $EpCAM^{+}$ /$CD45^{-}$ tumor cells was sorted and mRNA-expression analyzed. Despite a significant carryover of leukocytes, indicated by the detected CD45 mRNA in the samples (data not shown), it was decided to apply this method for patient material. Subsequently 17 patients' blood samples, including ten patients that had overt liver metastasis and three sample of mesenterial blood were prepared and sorted with FACS Aria II^{TM}. In each case between 20 and 100 cells were sorted as $EpCAM^{+}$ /$CD45^{-}$ by flow cytometry. Subsequent analysis by quantitative polymerase chain reaction (qPCR), however, revealed that these objects were probably unspecific fluorescent signals or no intact cells. B-actin expression was detected in twelve samples, however, late in qPCR (Ct > 20) and it correlated in most cases to CD45 expression indicating a carryover of leukocytes. Only one sample showed a putative EpCAM expression close to the detection limit of the qPCR (Ct > 37).

The micromanipulator enables the isolation of single CTCs

Since cell sorting by flow cytometry was not successful and bears the risk of losing the rare CTCs, it was decided to apply a micromanipulator for CTC isolation. After density centrifugation and the aforementioned anti-CD45 beads based pre-enrichment strategy CTCs can be identified in stained blood samples with the help of fluorescence microscopy by using EpCAM-directed antibodies. Subsequently, the micromanipulator that is mounted on the fluorescence microscope allows the direct pipetting of single cells with the help of a small glass capillary (Figure 11).

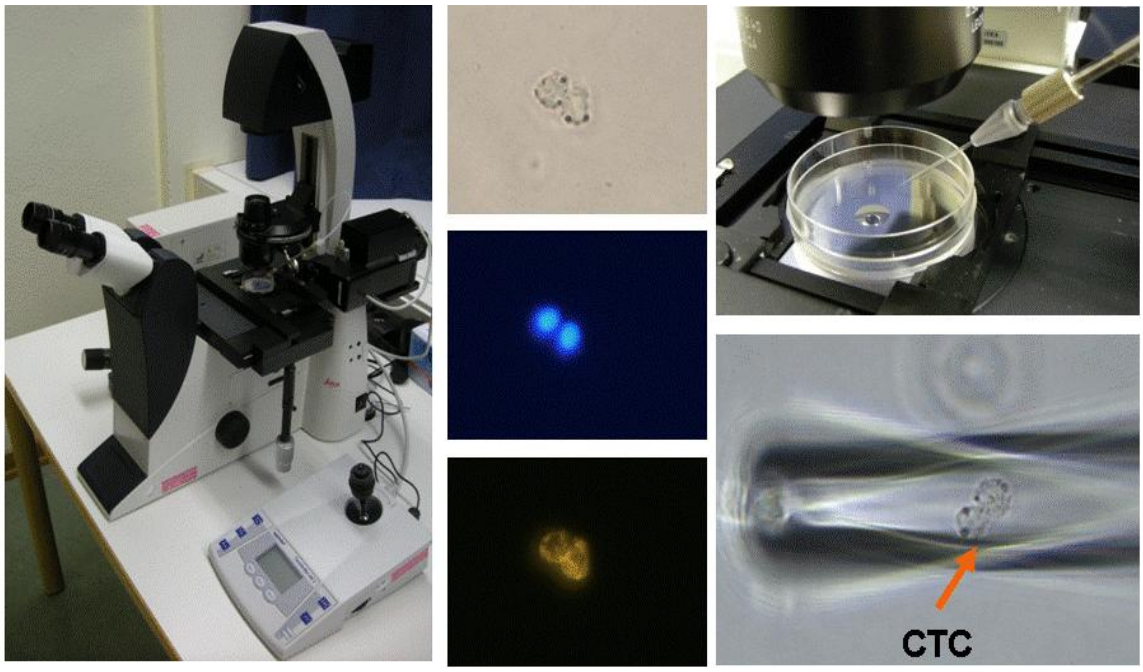


Figure 11: The micromanipulator. Cells in suspension are picked with a micromanipulator (A) mounted on a microscope. Fluorescence microscopy allows the identification of stained cells. (B=Bright field, C=Nuclear stain /DAPI, D=Cytokeratin-PE staining) Subsequently, cells can be aspirated with the help of an exchangeable glass capillary (E +F). The picking of two aggregated CTCs is shown.

A similar approach to isolate DTCs from bone marrow samples

To assess if the developed CTC isolation protocol can be applied for the isolation of DTCs from bone marrow specimens as well we analyzed the expression of EpCAM and CD45 in a bone marrow sample of a healthy donator (Figure 12 A). The majority of bone marrow cells was found to express the leukocyte marker CD45. Therefore, depletion of bone marrow cells with the help of anti-CD45-beads is suitable for DTC enrichment. Despite a small portion of the bone marrow cells seem to express EpCAM to a low extent ($EpCAM^{low}$) the identification of cancer cells spiked in the bone marrow samples is still possible because EpCAM expression in tumor cells ($EpCAM^{high}$) was much stronger than in the BMDCs. Flow cytometry analysis displayed the cancer cell population clearly apart from the bone marrow cells indicating that the strong EpCAM expression of cancer cells allows the visual identification of DTCs by fluorescence microscopy (Figure 12 B).

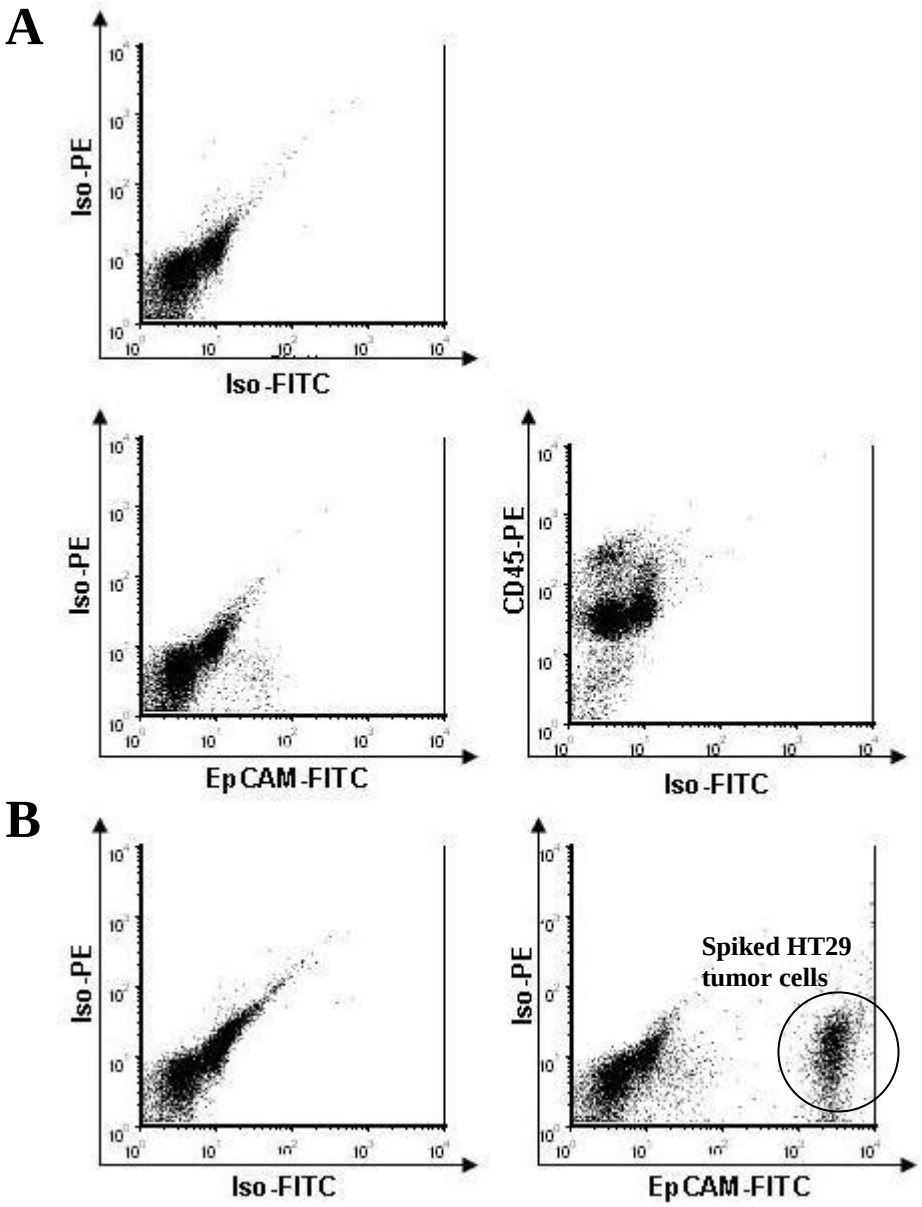


Figure 12: Surface antigen expression enables identification of DTCs: **A:** Bone marrow samples were analyzed for the expression of CD45 and EpCAM. **B:** A bone marrow sample was spiked with HT29 cancer cells. Tumor cells can be distinguished from bone marrow cells due to the high expression of EpCAM.

An mRNA expression study on CRC - associated CTCs and DTCs

We assume that tumor cell dissemination is associated with distinct changes in the mRNA expression profile. The present study ought to show if CTCs and DTCs display a specific transcriptional signature that differs from the corresponding cancer tissue.

Validation of the cDNA amplification strategy for the analysis of single cells

Despite remarkable progress during the last years, mRNA expression analysis of low cell numbers is still challenging and only possible if sophisticated amplification strategies are applied. To make sure that the intended amplification kit is suitable for the analysis of the low cell numbers ten and lesser cells were isolated for mRNA extraction. After cDNA amplification mRNA expression was assessed by quantitative RT-PCR. Figure 13 A shows that target gene mRNA expression of even single cells was detectable. For validation of the mRNA expression analysis ten tumor cells were isolated repeatedly in six independent experiments. The obtained mRNA was extracted and amplified. Subsequently, the relative expression of EpCAM, CK19 and CK18 normalized to β-actin was examined (Figure 13). Since cells were cultured under equal conditions, we expected low inter-experimental variations of the analyzed genes. Despite single outliers the results showed comparable expression levels indicating the reliability of the used amplification technique.

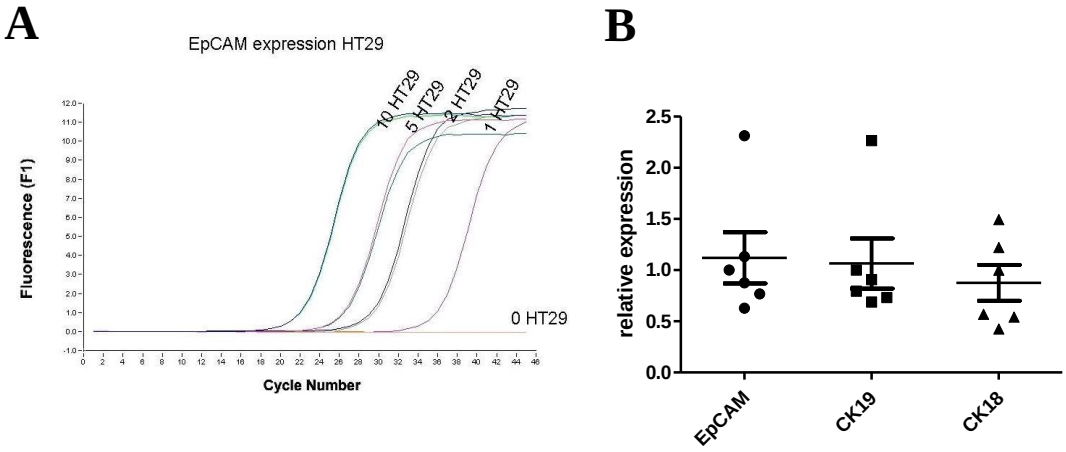


Figure 13: Validation of the cDNA amplification strategy. A RT-PCR analyses of low cell numbers. EpCAM expression of even a single HT 29 tumor cell was detectable. **B** Six independent target gene expression analyses of ten pooled tumor cells by RT-PCR normalized to β-actin. The tumor cells that were cultured under equal conditions show a comparable gene expression resulting in a low inter-experimental standard deviation.

Unambiguous mRNA expression profiles of CTCs and DTCs reveal biological and therapeutic implications

Between June 2010 and October 2011 72 blood samples and 9 bone marrow samples of CRC patients who underwent surgical resection for CRC in the department of surgery of the University Clinic Heidelberg were analyzed. Preferentially patients with overt liver metastases were selected to increase the chances to obtain blood samples containing CTCs. According to the results obtained from the CellSearch^{TM} system, that indicate a higher number of CTCs in the tumor draining veins, the blood was taken from a mesenteric vein or in case of liver metastases from the liver vein. CTCs were enriched and isolated as previously described by a manual enrichment protocol to enable the isolation of living cells. Figure 14 shows two images of CTCs as observed after EpCAM staining in a processed patient's blood sample. Additional blood samples were analyzed by the CellSearch^{TM} system as an independent control of the used enrichment and isolation method.

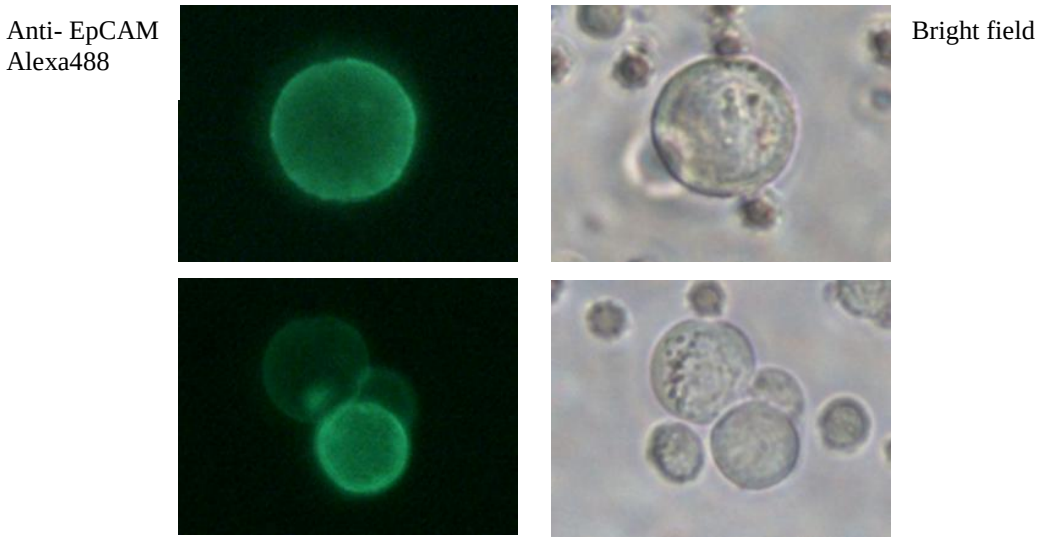


Figure 14: Human CTCs. CTC specimens from patient HD2095 stained with anti-EpCAM-Alexa488

At least one CTC was detected in 56 % (40) of the samples by the CellSearch^{TM} system in blood samples obtained either from the mesenteric or the liver vein. However, in only 14 % (10) of the cases it was possible to successfully enrich CTCs from the blood and in 22% (2) of cases to isolate DTCs out of bone marrow samples. Mostly, CTCs were found with the manual enrichment approach only when high numbers of CTCs were detected by the CellSearch^{TM} system in parallel. CTCs/DTCs appeared usually as single cells. Surprisingly, cell clusters (tumor microemboli) were only rarely observed probably due to the applied enrichment procedure. Subsequently to the detection and

isolation CTCs/DTCs of single patients were pooled for RNA extraction and amplification (Table 4).

Table 4: Patients characteristics. Overview about patients from whom CTCs or DTCs were obtained for mRNA expression analysis. No data about DTC incidence in bone marrow samples can be obtained from the CellSearch™ system (indicated by X) since the device does not allow the processing of bone marrow samples.

patient	sex	age	UICC stage	Blood sample (LV= liver vein, MV= mesenteric vein)	No. of CTCs detected by CellSearch (7.5 ml blood)	Manually isolated CTCs for mRNA expression analysis
CTC						
HD 2091	male	57	IV	LV	5	4
HD 2095	female	69	IV	LV	469	2 x 15
HD 2130	female	57	IV	LV	0	3
HD 2201	male	64	IV	LV	2	4
HD 2231	male	62	IV	LV	8	4
HD 2257	female	55	IV	LV	5	4
HD 2288	female	59	IV	LV	82	7
HD 2334	male	73	III	MV	208	3 x 8
HD 2340	male	62	III	MV	3	2
DTC						
HD 2401	male	39	II	bone marrow	X	16
HD 2417	male	70	IV	bone marrow	X	5

Although mRNA expression analysis of single cells would be feasible, we decided to pool CTC populations to obtain more valid data and to increase the chances to detect even lowly expressed genes. Robust signals in qPCR were obtained when at least four cells were pooled for analysis. In addition, a similar number of cells from matched tissue of either primary tumor or liver metastases was taken to compare mRNA expression. Results are depicted in heat maps (Figures 15 – 18). A large-scale expression screen comprising 47 genes related to CTC/DTC identification, common CRC biomarkers, EMT, stemness and invasiveness were selected to describe CTC/DTC characteristics. In addition, we included CD45, a pan leukocyte marker that is not expressed on epithelial cells, in the study. Although it was a major aim to analyze CTC/DTC specimen without any disturbing background of leukocyte contaminations, it seemed to be impossible to obtain samples that were free of any CD45 signal in the qPCR. The ubiquitary presence of free leukocyte-derived CD45 mRNA in the patients' samples probably leads to an unavoidable detection of CD45 mRNA in qPCR. Since we detected CD45 always in a comparable range among CTCs/DTCs and cancer tissue

samples, the similar CD45 signal in qPCR should be rather considered as technique-derived noise. However, leukocyte carryover can not be excluded in general. Therefore, the supposed impurity of the specimens requires a cautious evaluation of the obtained expression profiles. In particular the expression of genes that are significantly expressed in the control leukocyte and bone marrow samples including CD44, Vimentin, CXCR4 and CD47 are difficult to assess. The robust signal of those genes in qPCR analysis indicates that they are expressed on tumor cells, however, it is difficult to state about their up or down-regulation during the dissemination process.

CTC expression profiles reflect a fairly heterogeneous picture

To compare general mRNA expression levels of CTCs and matched cancer tissue samples the Ct values obtained from qPCR data are depicted as heat map in figure 15. Since carryover of leukocytes might have introduced bias in the results; the mRNA expression profiles of three control leukocyte samples are shown in addition. The mRNA expression profile of CTCs is unambiguously distinct form leukocytes. The expression of EpCAM, CK18, CK19 and CEA in the CTCs and the absence of these markers in control leukocyte samples clearly confirmed the epithelial origin of the CTC specimens. EpCAM, CK18 and CK19 appeared to be well suited for CTC identification. CK20, however, seemed rather unstably expressed in CTCs. In addition, Ki67 was found to be absent or only lowly expressed in CTCs, which might indicate a proliferative arrest in CTCs. In general, gene expression of CTCs was found to be very heterogeneous. Unsupervised hierarchical clustering analyses using the Spearman rank correlation did not show a correlation among CTCs or among tumor and metastases and no matched correlation between CTCs and the corresponding cancer tissues (data not shown).

With the present approach we were only able to evaluate the expression of a small panel of genes. Within this panel common traits of CTC samples seem to be limited to a few epithelial markers. Other genes seem to be expressed only occasionally in individual CTC specimens. For instance, the expression of stem cell markers like Bmi-1 and Ascl2 was limited to CTCs of HD2201 or HD2095 respectively. The same applies for the pro-apoptotic protein Bax that was detected in CTCs of patients HD2095 and HD2334 but not in other CTC specimens. Therefore, apoptosis seems to be attributed only to some CTCs, however, not a general feature.

Interestingly, it was not possible to confirm an enhanced expression of EMT-associated genes in CTCs. Although Vimentin seems to be expressed in CTCs, EMT-related transcription factors Snail and Twist as well as hFN, N-cadherin were mostly not detectable.

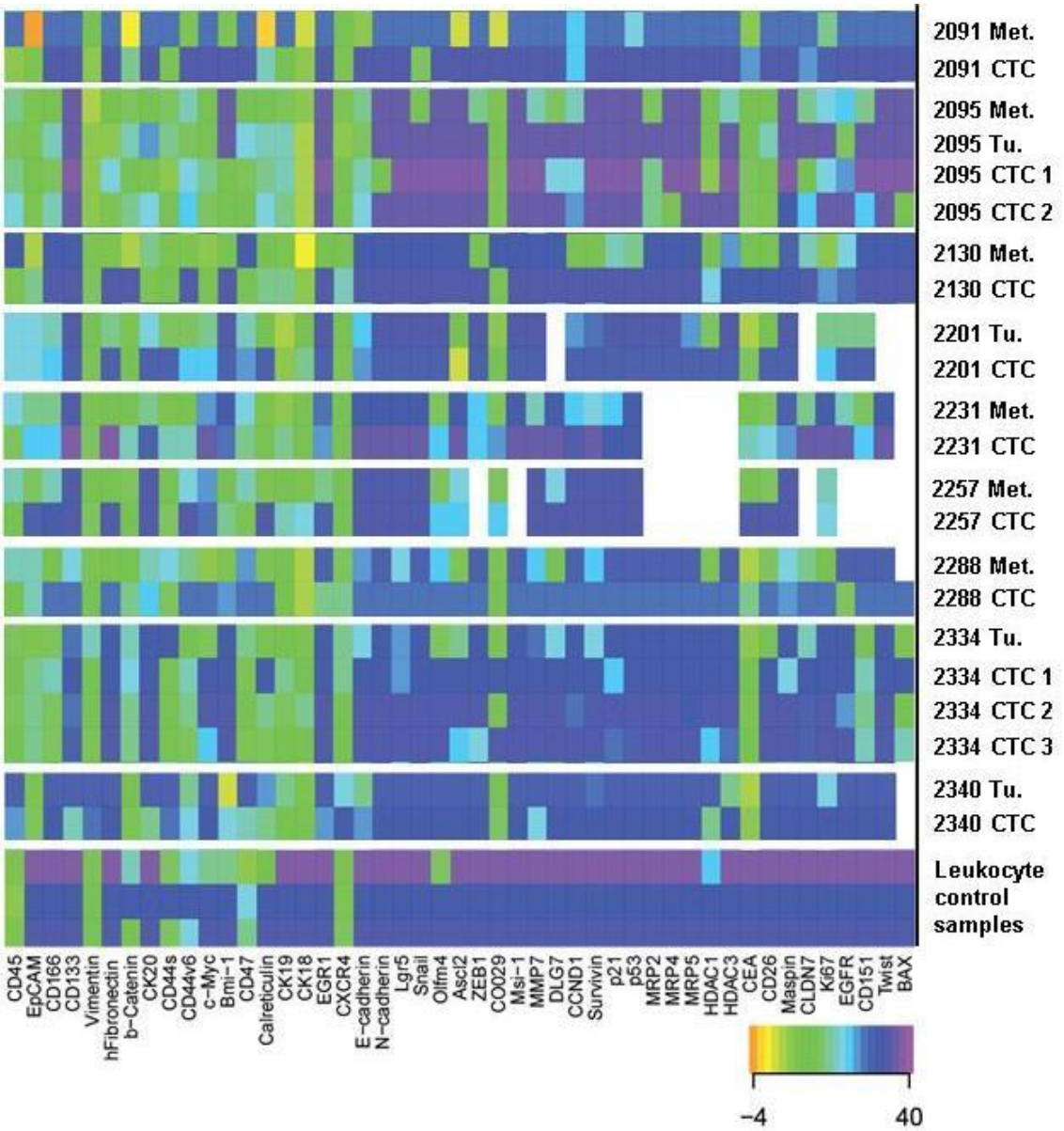


Figure 15: mRNA expression of CTCs. mRNA expression analysis of isolated CTC samples and matched samples of tumor (Tu.) or metastasis (Met.). Additionally, three leukocyte samples were analyzed to evaluate the influence of possible background carryover. Color scale represents ΔCt values of target and house keeping genes (β-actin). Low ΔCt values (yellow and green) indicate highly expressed genes. High ΔCt values (blue and purple) indicate lowly expressed genes. A white area indicates cases without available data.

CTCs rather resemble metastases than primary tumors in gene expression

To compare gene expression between CTCs and the matched tissue samples two further heat maps are shown in figure 16. The representation of ΔΔCt values as color code facilitates the identification of differentially regulated genes among CTCs and matched cancer tissue samples. Interestingly, evaluated genes seem to be rather up-regulated in CTCs when compared to matched tumor samples, however, rather down-regulated or equally expressed when compared to liver metastases. For instance CD47, which was reported to be involved in immune escape [105], was found up-regulated in CTCs only when compared to tumor samples, however, equally expressed referring to metastases. Or the expression of the hyaluronic acid receptor and stem cell marker CD44s which seemed to be up-regulated in CTCs compared to primary tumor samples, but rather similar expressed referring to matched metastases. Other evaluated stem cell markers (CD166, Bmi-1, Lgr5, ZEB-1, and DLG7) were only rarely detected in CTCs and in the corresponding samples of tumor or metastasis. For CD133 no significant differential regulation was observed.

With regard to EMT CTCs retained Vimentin expression after dissemination. If compared to tumor samples it seems to be even up-regulated. However, as previously mentioned Vimentin is expressed in leukocytes as well and since leukocyte carryover can not be excluded an up-regulation of Vimentin in CTCs can only be assumed. With further regard to EMT we were not able to show an up-regulation of mesenchymal genes like N-cadherin, Snail or Twist in the CTCs. Fibronectin expression seemed even to be lower expressed in CTCs than in matched metastasis samples. E-cadherin was found to be slightly down-regulated in CTC samples compared to the primary tumors; however, in comparison to metastases its expression was rather similar. The proliferation marker Ki67 seemed to be rather equally expressed among CTCs and cancer tissues. Genes attributed to apoptosis and survival, including p21, p53, Survivin and BAX, were not found to be differentially regulated. A significant up-regulation of metastasis-associated genes like CD44v6, CD26 or CO029 in CTCs was not observed.

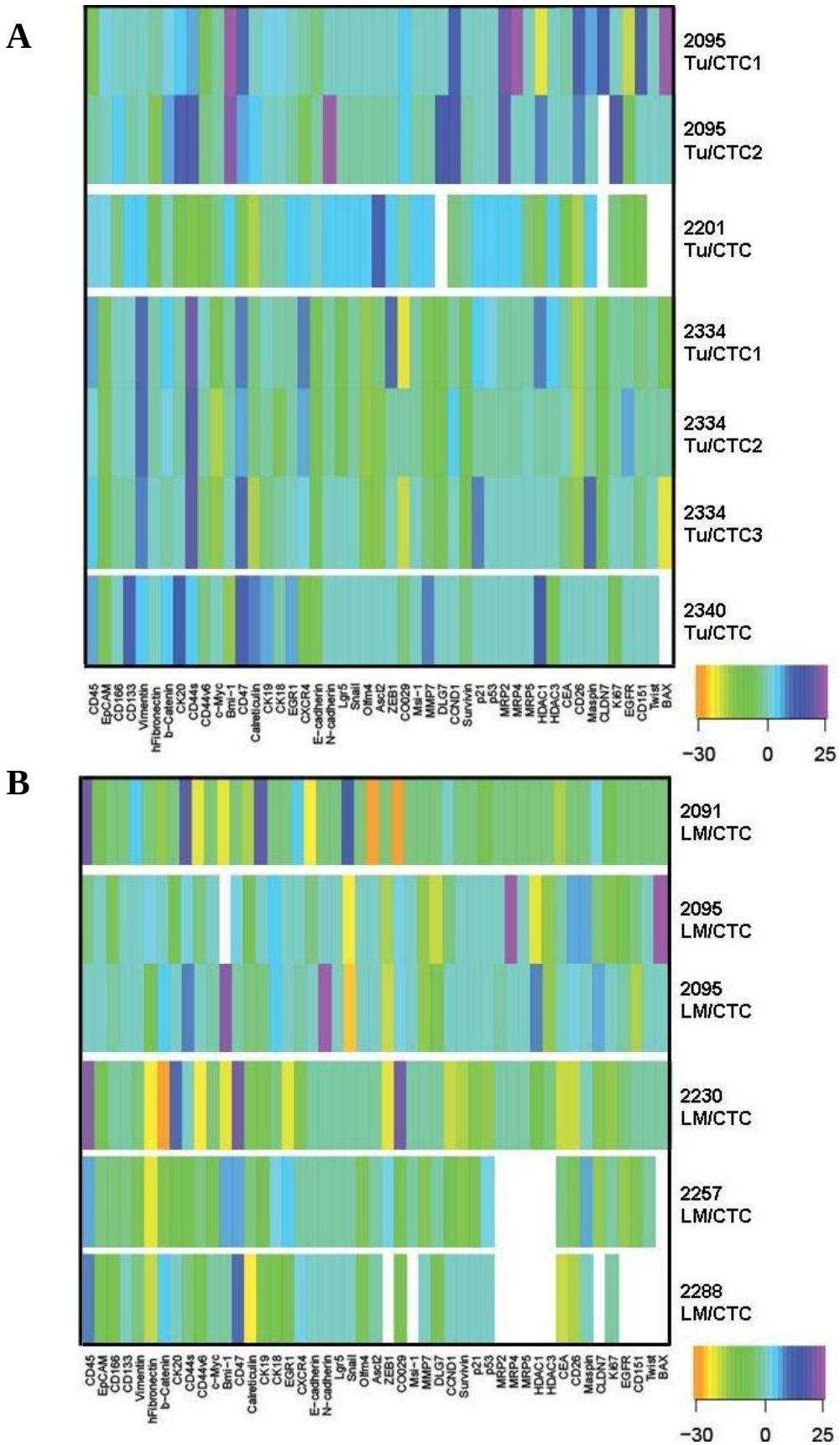


Figure 16: Compared mRNA expression of CTCs and matched tumor (A) or liver metastasis (B). ΔΔCt values of ΔCt tumor/metastasis – ΔCt CTCs are depicted as heat map. Color scale indicates highly expressed genes of cancer tissue in yellow and green. Similar expression in both compartments is shown in faint green. Genes that are highly expressed in CTCs are depicted in blue and purple. A white area indicates cases without available data.

DTCs are probably dormant and lack EGFR expression

Unfortunately, only mRNA expression data of two samples of DTCs is available (Figures 17 and 18). Despite this small number of specimens, a first trend seems to be evident. Unlike the CTCs the DTCs have lost CK19 expression in the bone marrow. Only CK18 and EpCAM expression was reliably detected in DTCs and confirmed the epithelial origin. E-cadherin appeared to be down-regulated. However, an up-regulation of EMT-associated genes was not detected. Analyses indicated that DTCs were probably non-proliferative and dormant as gene expression seemed to be in general down-regulated irrespectively if compared to tumor of metastasis (Figure 18). Numerous genes detected in tumor or metastases samples were not expressed in DTCs including the proliferation marker Ki67 as well as CD151, CEA, CD26, CO029 and EGFR. A significant expression of several CRC-associated stem cell markers, such as Lgr5, Ascl2 or ZEB1 was also not detectable. Olfm4 and DLG7 seem to be expressed in DTCs. However, these genes are expressed in control bone marrow samples as well. Since undesired carryover can not be excluded, expression of Olfm4 and DLG7 requires independent confirmation.

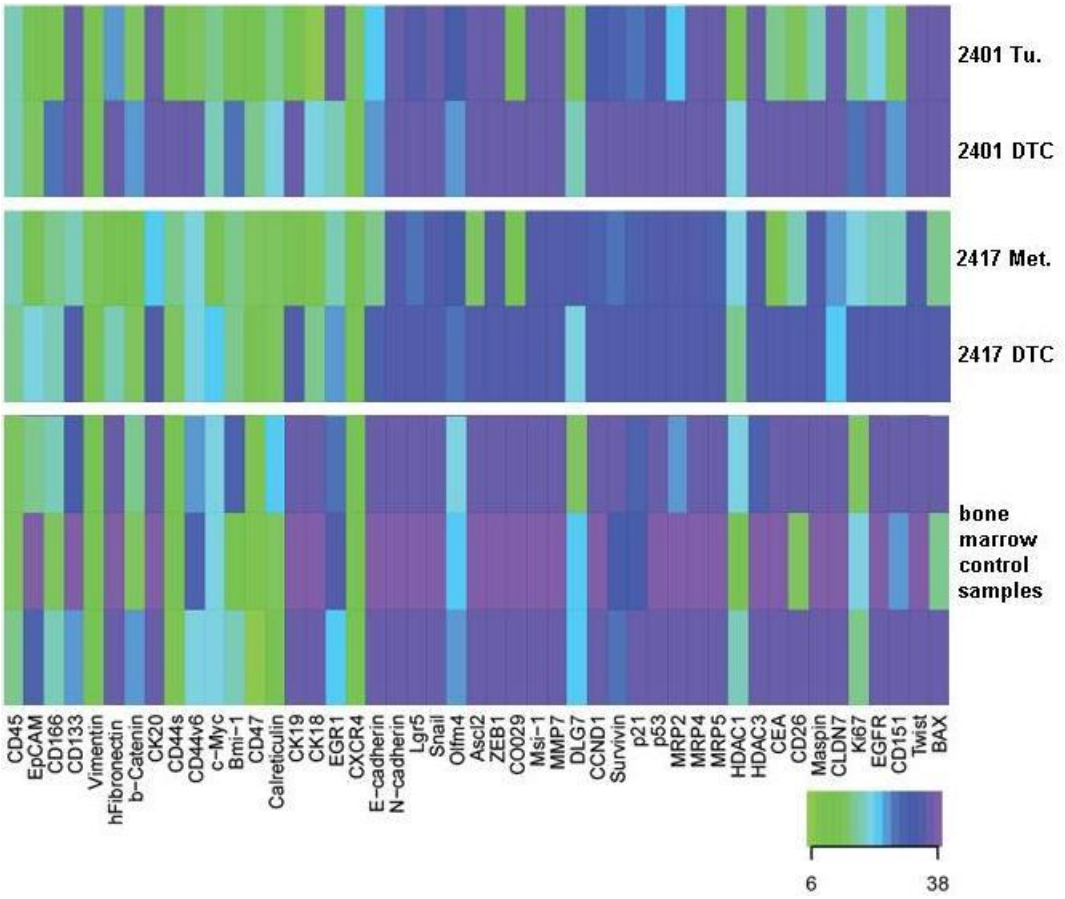


Figure 17: mRNA expression of DTCs. mRNA expression analysis of isolated DTC and matched samples of tumor (Tu.) or metastasis (Met.). Additionally, three bone marrow samples were analyzed to evaluate the influence of possible background carryover. Color scale represents ΔCt values of target and house keeping genes (β-actin). Low ΔCt values (green) indicate highly expressed genes. High ΔCt values (blue and puple) indicate lowly expressed genes.

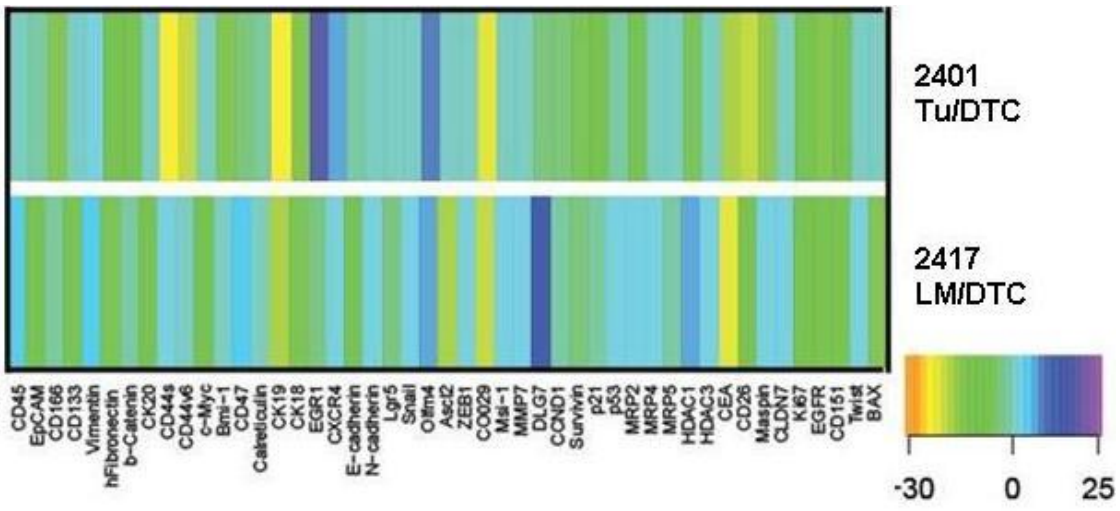


Figure 18: Compared mRNA expression of DTCs and matched cancer tissue. ΔΔCt values of ΔCt tumor/metastasis – ΔCt DTCs are depicted as heat map. Color scale indicates highly expressed genes of cancer tissue in yellow and green. Similar expression is shown in faint green. Genes that are highly expressed in DTCs are depicted in blue and purple.

Genomic characterization of CTCs

Today's therapeutic approaches against cancer come away from just targeting rapid growing cells and try to interfere with specific pathways that are genetically altered in tumor cells. However, genetic and epigenetic heterogeneity is common in human cancers [106, 107]. Thus, the need to determine the molecular profile of cancer cells to predict therapy response is becoming increasingly important. With this study we want to assess the possibilities to characterize single CRC-associated CTCs on the genomic level in the strict sense of their mutational profile.

CTCs carry typical point mutations in CRC-associated genes

CRC-associated CTCs that were enriched and detected by the CellSearch™ system were isolated as single cells with a micromanipulator to analyze their mutational profile. According to the CellSearch™ guidelines CTCs were defined as $EpCAM^+$ / CK^+ (CK8, CK18, CK19 or in combination)/ $DAPI^+$ / $CD45^-$. A specific method to amplify gDNA of single cells published by Klein and colleagues [108] was applied to yield sufficient analyzable material. We were interested in common CRC-associated mutations including *BRAF* V600E, mutations in codons 12 and 13 in the *KRAS* gene, and aberrations within the DNA binding domain of *TP53* (Figure 19).

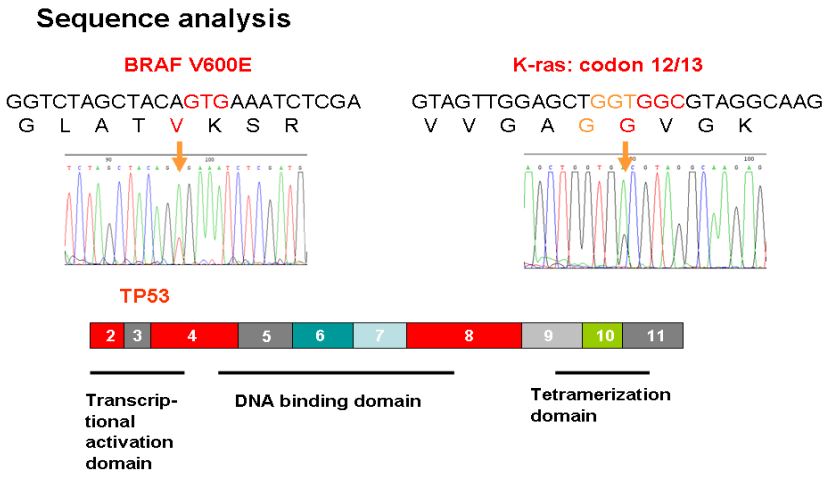


Figure 19: Mutation analysis. Single CTCs were analyzed for mutations in *BRAF* (V600E), *Kras* (codon 12 and 13), and in the DNA binding domain of *TP53* (exons 4-8) by direct sequencing. The figure depicts the included genomic regions.

It was possible to isolate and analyze between 1 and 14 CTCs from 31 individual patients (in total 126 CTCs) (supplementary figure 1). 68.8% of included patients were diagnosed as UICC stage IV, 21.9% UICC III, 3.1% UICC II, 6.2% UICC I (supplementary figure 2). In CTCs of 16 patients no mutations were detected within the evaluated genomic regions of interest. A mutation of codon 12 or 13 of the *Kras* gene was detected in CTCs of six patients (19.4%). The V600E mutation of *BRAF* was found in a CTC of one patient (3.2%). CTCs of eleven patients were found to carry a mutation in the DNA binding domain of *TP53* (35.5%) (Table 5).

Table 5: Overview about detected point mutations in CTCs.

	Patients without a detected mutation	Kras C12/C13	Braf V600E	DNA-BD TP53
No. of patients with indicated mutations (31 patients, 126 CTCs, 1-14 CTCs per patient)	16 (51.6%)	6 (19.4%)	1 (3.2%)	11 (35.5%)

Identification of CTCs with MSI within coding and non-coding genomic regions

MSI is a common phenomenon in CRC [12]. Currently, conflicting data exist about the relevance of MSI for prognosis and therapy response in CRC [109, 110]. We wanted to assess if the detection of MSI in amplified gDNA specimens of CTCs is possible. Since our single cell gDNA amplification approach depends on the sequence specific nuclease digestion by MSE I, it was not possible to include the same panel of MS markers that is recommended by the national cancer institute (NCI) (Bethesda guidelines) [111]. Therefore, we typed CTCs at the three mononucleotide MS markers *NR21, NR24* and *BAT 25* that are widely accepted to indicate overall MSI in a cell [112].

At least to our knowledge studies on MSI in amplified genomic specimens of single CRC-associated CTCs have not been reported until today. To exclude false positive typing, what might be caused by artifacts occurring during gDNA amplification, we included 35 normal leukocytes processed similarly to the CTC samples in the study. None of the analyzed leukocytes showed any signs of MSI within the evaluated MS.

CTCs showing MSI were found in the blood samples of two patients (HD2215 and HD2341; 6.5%) (Figure 20). Those samples were subsequently analyzed for MSI in MS within coding genomic regions, so called coding microsatellites (cMS). The panel of cMS included Bax, DD5, FLT3LG and ELAVL3, because the mutation frequency of these MS-containing genes is well-known [113] and the applied gDNA amplification method did not interfere with their evaluation. The three analyzed CTCs of HD2341

showed MSI in at least two MS including *NR21* and the cMS within *ELAVL3* (Table 6). *ELAVL3* codes for a RNA binding protein of the HuR family and is involved in the posttranscriptional regulation of mRNA expression. According to the recommendations of the NCI tumors with mutations in two of a panel of five MS are considered to have high MSI (MSI-H). If only one of the five MS is altered, the tumor is considered to have low MSI (MSI-L). Although it was not possible to evaluate the same set of MSI markers as suggested by the NCI, we typed the tumor cells of patient HD2341 as MSI-H because of two altered MS. In case of patient HD2215 two of nine typed CTCs showed MSI in *NR21*; however, no further MSI in any of the other analyzed markers was detected. Therefore, the aberration in only one MS marker refers to a MSI-L tumor. MSI-L was reported for various tumors including colorectal, breast, endometrial and ovarian cancers [114, 115] and might not be attributed to defects in cellular DNA mismatch repair systems. Instead, some DNA replication errors that occur stochastically and to a normal rate might have remained unrepaired due to an overload of the mismatch repair system as caused by high rate cell divisions in cancer.

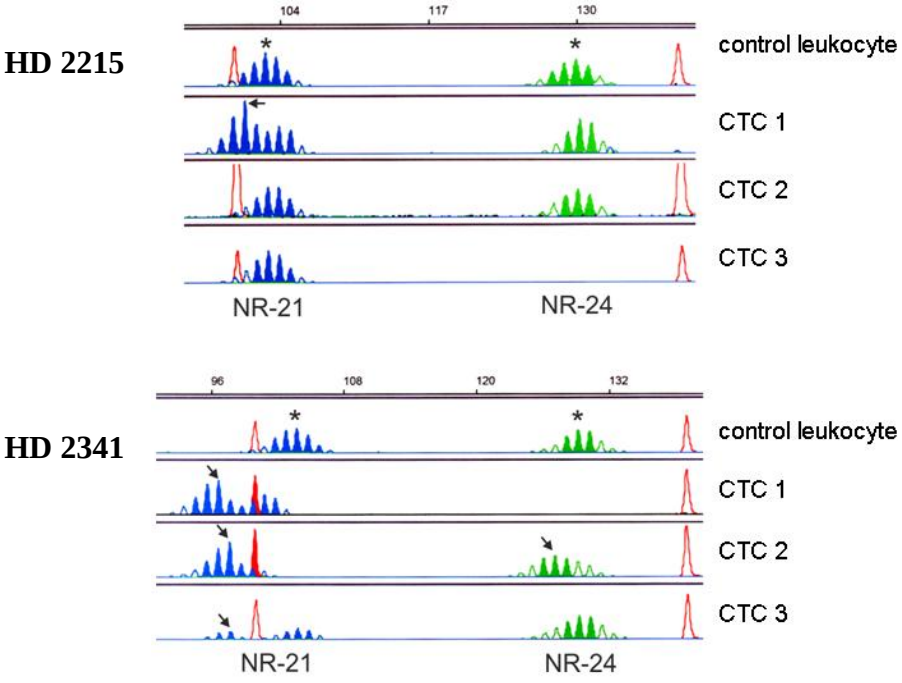


Figure 20: MSI in CTCs. Defects in DNA repair mechanisms can lead to lengthening or shortening of microsatellites (MS). Amplified length polymorphism (AFLP) analyses of CTCs and normal control leukocytes of patients HD2215 and HD2341 for NR21 (blue) and NR24 (green) are shown. Stutter bands arise since polymerase chain reaction amplification creates DNA fragments that are one or several repeats longer or shorter than the actual allele. Therefore, the highest detected peak was determined as the reference peak. Red peaks represent internal length control markers. Asterisks indicate the unchanged MS-length as detected in control leukocytes. Arrows indicate an altered MS length in single CTCs. For HD2215 CTC3 and HD2341 CTC1 no signal for NR24 was detected. This might be caused by losses of gDNA fragments during the single cell amplification procedure.

Table 6: cMS analysis of MSI$^+$-CTCs. Three CTCs and one leukocyte (PBMC) of patient HD 2341 were analyzed for microsatellite instability (MSI) within coding DNA regions. No data was available for microsatellite sequence within BAX of CTC 2341 mes1 and 2341 PBMC1. MSI = microsatellite instability, MSS = microsatellite stable

CTC	NR21	NR24	BAX	DD5	FLT3LG	ELAVL
2341mes1	**MSI**	-	-	MSS	MSS	**MSI**
2341mes2	**MSI**	**MSI**	MSS	MSS	MSS	**MSI**
2341mes3	**MSI**	MSS	MSS	MSS	MSS	**MSI**
2341 PBMC1	MSS	MSS	-	MSS	MSS	MSS

Inter - CTC genomic heterogeneity confirms clonal diversity

In case of 16 patients more than one CTC could be analyzed. Six of those patients had a population of CTCs that were genomic heterogeneous (37.5%) with regard to mutations in *Kras*, *BRAF*, *TP53* or MSI (supplementary figure 2). For instance, we have analyzed nine CTCs from the patient HD2215 for MSI (Table 7). Two CTCs were obtained from a blood sample of the liver vein (LV-CTC) and seven were isolated from mesenteric venous blood compartment (MV-CTC). Seven CTCs did not show any signs of MSI in the evaluated markers but two MV-CTCs showed MSI indicated by a changed allele length of *NR21*. Additionally, a *Kras* mutation was detected in four MS stable (MSS) CTCs of that patient whereas the mutation was not found in the MSI-CTCs. Another case of inter–CTC cytogenomic heterogeneity was found for patient HD2328 (Table 8). CTCs were isolated from blood samples obtained from the mesenteric vein and the liver vein. Astonishingly, one half of the analyzed CTCs (3/6) isolated from blood of the liver vein carried a mutation in exon 5 of the *TP53* gene whereas the residual half did not. Intriguingly, we could not detect this mutation in CTCs obtained from the mesenteric blood compartment.

In summary, it was possible to detect MSI and CRC-associated genomic mutations in single CTCs. In addition, we found cases of inter-CTC genomic heterogeneity in CTC fractions of single patients. The presence of various tumor subclones within a patient is reflected by different CTC populations that may have derived from different metastatic sites.

Table 7: Genomic inter-CTC heterogeneity of CTCs from patient HD 2215.
CTCs and cancer tissue samples of patient HD 2215 were analyzed for microsatellite instability (MSI) and mutations within the indicated genomic regions. MSS = microsatellite stable, wt = wild type allele detected, n.d. = no data available.

Patient sample /CTC	Microsatellite stability	Kras C12/13	Braf V600E	TP53 E5/6	TP53 E7	TP53 E8
2215 Tumor	MSS	**G12S**	wt	wt	wt	wt
2215 liver met.	MSS	**G12S**	wt	wt	wt	wt
CTC LV1	MSS	**G12S**	wt	n.d.	n.d.	wt
CTC LV2	MSS	n.d.	n.d.	n.d.	n.d.	wt
CTC mes 1	MSS	n.d.	n.d.	n.d.	wt	n.d.
CTC mes 2	MSS	**G12S**	wt	n.d.	wt	wt
CTC mes 3	**MSI**	wt	n.d.	n.d.	n.d.	n.d.
CTC mes 4	MSS	**G12S**	wt	n.d.	n.d.	n.d.
CTC mes 5	MSS	wt	wt	wt	wt	wt
CTC mes 7	**MSI**	n.d.	n.d.	n.d.	n.d.	n.d.
CTC mes 8	MSS	n.d.	wt	n.d.	n.d.	n.d.

Table 8: Genomic inter-CTC heterogeneity of CTCs from patient HD 2328.
CTCs and a liver metastasis sample of patient HD 2328 were analyzed for microsatellite instability (MSI) and mutations within the indicated genomic regions. MSS = microsatellite stable, wt = wild type allele detected, n.d. = no data available.

Patient sample /CTC	Microsatellite stability	Kras C12/13	Braf V600E	TP53 E5/6	TP53 E7	TP53 E8
2328 liver met.	MSS	wt	wt	**F134C**	wt	wt
CTC LV1	MSS	wt	wt	**F134C**	wt	wt
CTC LV2	MSS	wt	wt	**F134C**	wt	n.d.
CTC LV3	MSS	wt	wt	wt	wt	wt
CTC LV4	MSS	wt	wt	wt	wt	wt
CTC LV5	MSS	wt	n.d.	n.d.	wt	wt
CTC LV6	MSS	wt	wt	wt	wt	wt
CTC LV7	MSS	wt	n.d.	**F134C**	n.d.	wt
CTC LV8	MSS	wt	wt	wt	wt	wt
CTC mes 1	MSS	wt	wt	n.d.	n.d.	wt
CTC mes 2	MSS	wt	wt	wt	wt	wt
CTC mes 3	MSS	wt	wt	wt	wt	wt
CTC mes 4	MSS	n.d.	n.d.	n.d.	n.d.	n.d.
CTC mes 6	MSS	wt	wt	wt	wt	wt
CTC mes 7	MSS	wt	wt	wt	wt	wt

Malignant identity and chromosomal profile of single CRC-derived CTCs can be confirmed by matrix-CGH

Chromosomal instability is a hallmark of CRC [11] and chromosomal aberrations were found to be of prognostic relevance [116, 117]. However, if specific chromosomal alterations exist that are significant for the progression from invasive carcinoma to metastatic disease is still a matter of debate. We aimed to evaluate in a pilot experiment if the amplified gDNA of single CTCs is suited for a comparative genomic hybridization (CGH) analysis with regard to compare the genomic profiles among CTCs and solid cancer tissue. Therefore, array-CGH analyses of provided gDNA samples were performed in the lab of Prof. Dr. N. Stöcklein at the University Clinic Düsseldorf. To proof if pre-amplified gDNA from single cells allows the detection of genomic aberrations, amplified gDNA from two leukocytes were hybridized against each other on a 60k array (Figure 21). As expected, the ratio profiles are balanced for all autosomes. The genomic imbalance found for the sex chromosomes correctly reflects the different sexes of the two hybridized samples.

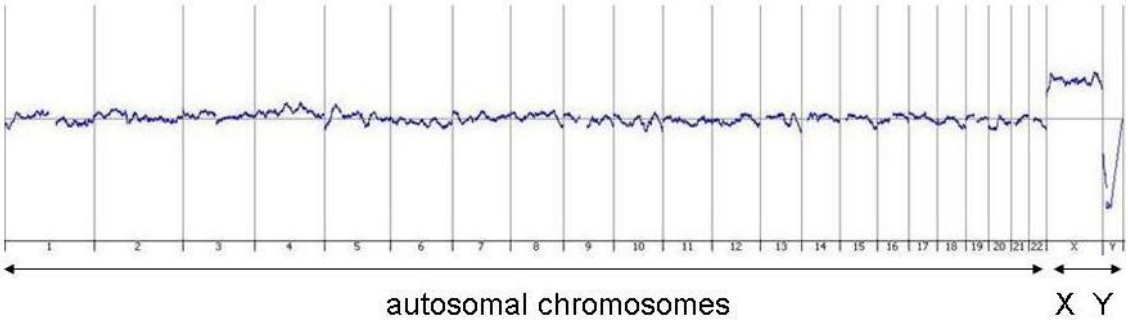


Figure 21: Validation experiment. Amplified gDNA of single male and female leukocytes were hybridized against each other. The balanced ratio profile found for the autosomes confirms unbiased gDNA amplification. The genomic imbalance detected for the sex chromosomes (X, Y) correctly reflects the different sexes (male and female) of the samples.

We analyzed in this pilot study the genomic profile of ten CTC samples with the help of a 180 k matrix-CGH array (supplementary figure 3). Three CTC samples were detected with a balanced CGH profile. Seven CTCs showed chromosomal aberrations that are typically found in CRC (Figure 24) (compare data base of University of Zurich: www.progenetix.org). These genomic aberrations comprised chromosomal gains in chromosomes 7, 8q, 13, 20 and losses in chromosomes 4, 8p, 18. Figure 24 depicts the penetrance plot of pooled data from all analyzed CTC specimens. The pilot study confirms that the evaluation of chromosomal instability in single CTCs is possible.

Comparing genomic profiles of CTCs to matched cancer tissue

Today, individual therapy decisions rely on cancer tissue biopsies [118]. To assess if the genomic profile of CTCs should be taken into consideration as well, we aimed to evaluate if the genomic profile of CTCs might differ from matched cancer tissue specimens.

Patterns of mutations often differ among CTCs and matched cancer tissue

The patterns of mutation of CTCs were compared with that of macro-dissections of the respective primary tumor or liver metastasis (Supplementary figure 1). For 26 of 31 analyzed patients at least one of the included mutations was detected either in CTCs or in the corresponding malignant tissue. In 14 of 26 cases the detected mutations matched between the tissue and at least one analyzed CTC. A single time MSI analysis did not match between CTCs and malignant tissue. Sequencing results were not corresponding to each other eight times for *Kras*, three times for *BRAF* and three times for the sequenced region in *TP53*.

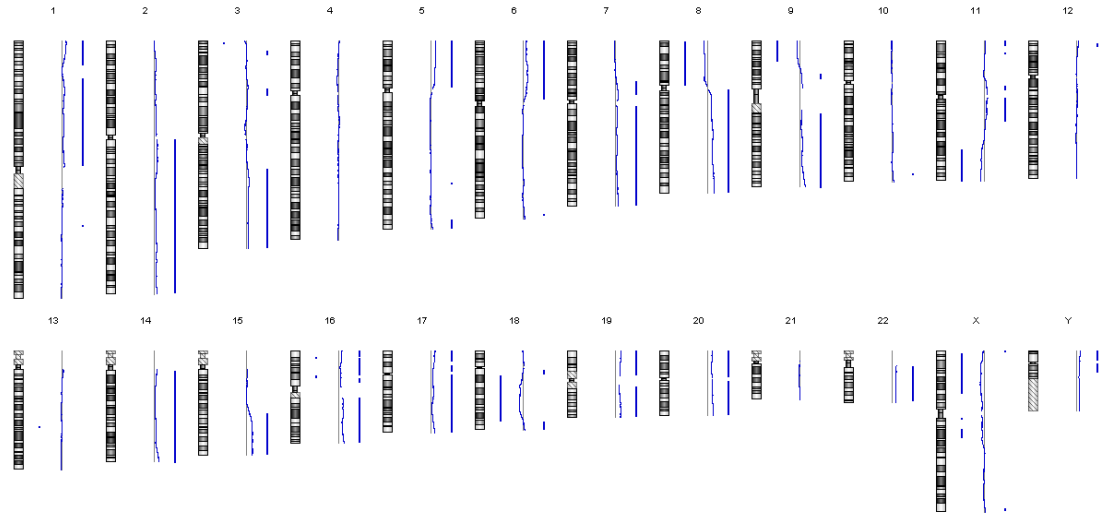
CGH profiles of CTCs reflect individual clones of matched cancer tissue

Additionally, array-CGH profiles of matched macro-dissected tumor or metastasis tissues could be generated for eight of the ten analyzed CTC samples (Supplementary figure 3). Seven tissue samples displayed typical CRC-associated genomic aberrations, one sample showed a balanced CGH profile without significant chromosomal changes. Twice the CGH profiles matched largely among the analyzed CTC and the matched cancer tissue. In six cases the CGH profiles appeared to be rather disparate between CTC and the matched cancer tissue sample.

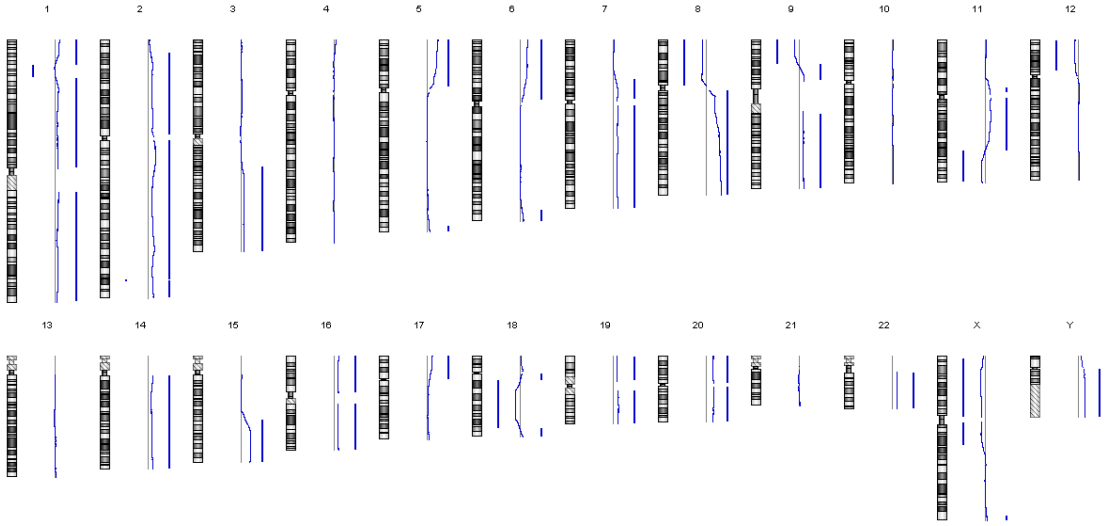
For instance, the CGH profiles a single liver vein - derived CTC and macro-dissected tissue from a liver metastasis of patient HD2095 showed a broadly similar chromosomal profile (Figure 22). Despite some disparate aberrations (chromosome 1q, 2p and 17q), lots of identical chromosomal gains and losses indicate a common clonal origin of the analyzed specimens. Disparity might be attributed to the use of macro-dissected tissue from the liver metastasis that represents the average of chromosomal aberrations found in the metastatic lesion.

For liver a vein –derived CTC and liver metastasis of patient HD2288 we found largely disparate array-CGH profiles. We assume that the analyzed CTC does not reflect the prevailing clone of the metastasis (Figure 23). Figure 24 compares the accumulative

CGH penetrance plots of pooled CTC samples with pooled data obtained from array-CGH analysis of tumor or metastasis tissue. Typical genomic aberrations associated with CRC were found. This includes chromosomal gains at chromosomes 7, 13, 20q and losses at chromosomes 4, 8p, 17p and 18q.

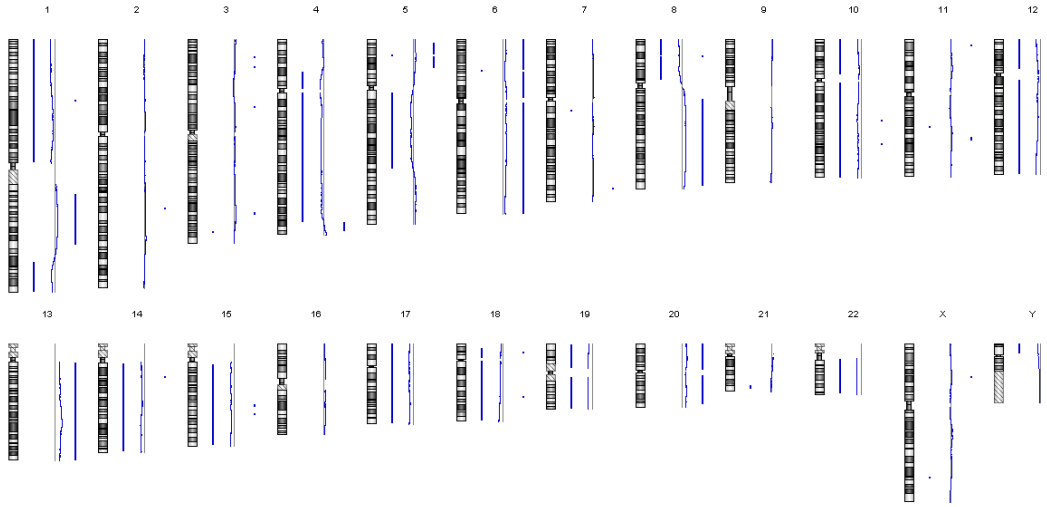


HD 2095 CTC

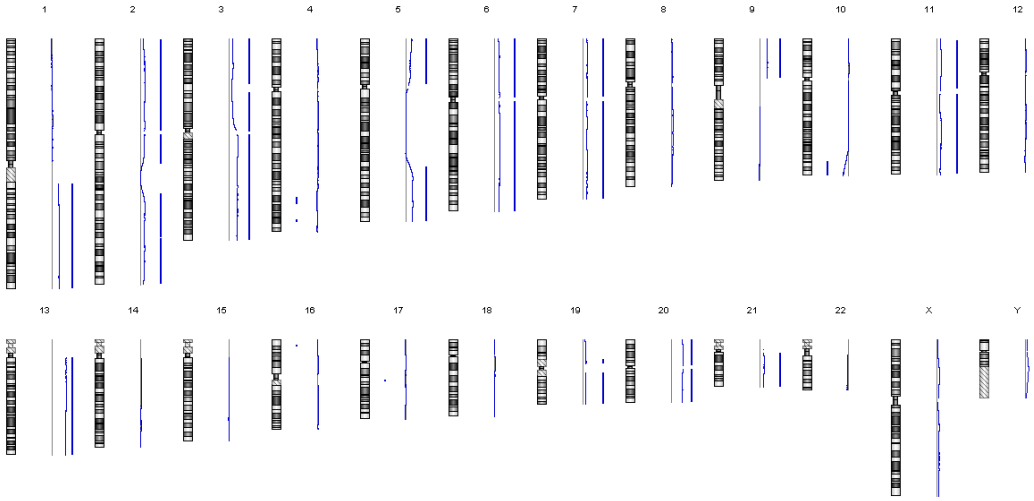


HD 2095 liver metastasis tissue

Figure 22: Array-CGH profiles of a liver vein - derived CTC and liver metastasis tissue of patient HD 2095. The human set of chromosomes is shown (1-22, X, Y). Beside each chromosome the detected signal intensity of the array-CGH analysis is depicted. Amplitudes to the left reflect chromosomal losses. Amplitudes to the right reflect chromosomal gains. The detected chromosomal aberrations are broadly similar among CTC and metastasis tissue and might indicate a common clonal origin of the analyzed specimens.



HD 2288 CTC



HD 2288 liver metastasis tissue

Figure 23: Array-CGH profiles of a CTC and liver metastasis of patient HD 2288. The human set of chromosomes is shown (1-22, X, Y). Beside each chromosome the detected signal intensity of the array-CGH analysis is depicted. Amplitudes to the left reflect chromosomal losses. Amplitudes to the right reflect chromosomal gains. The detected chromosomal aberrations are largely disparate among CTC and metastasis tissue and might indicate genetic clonal diversity of the analyzed cancer.

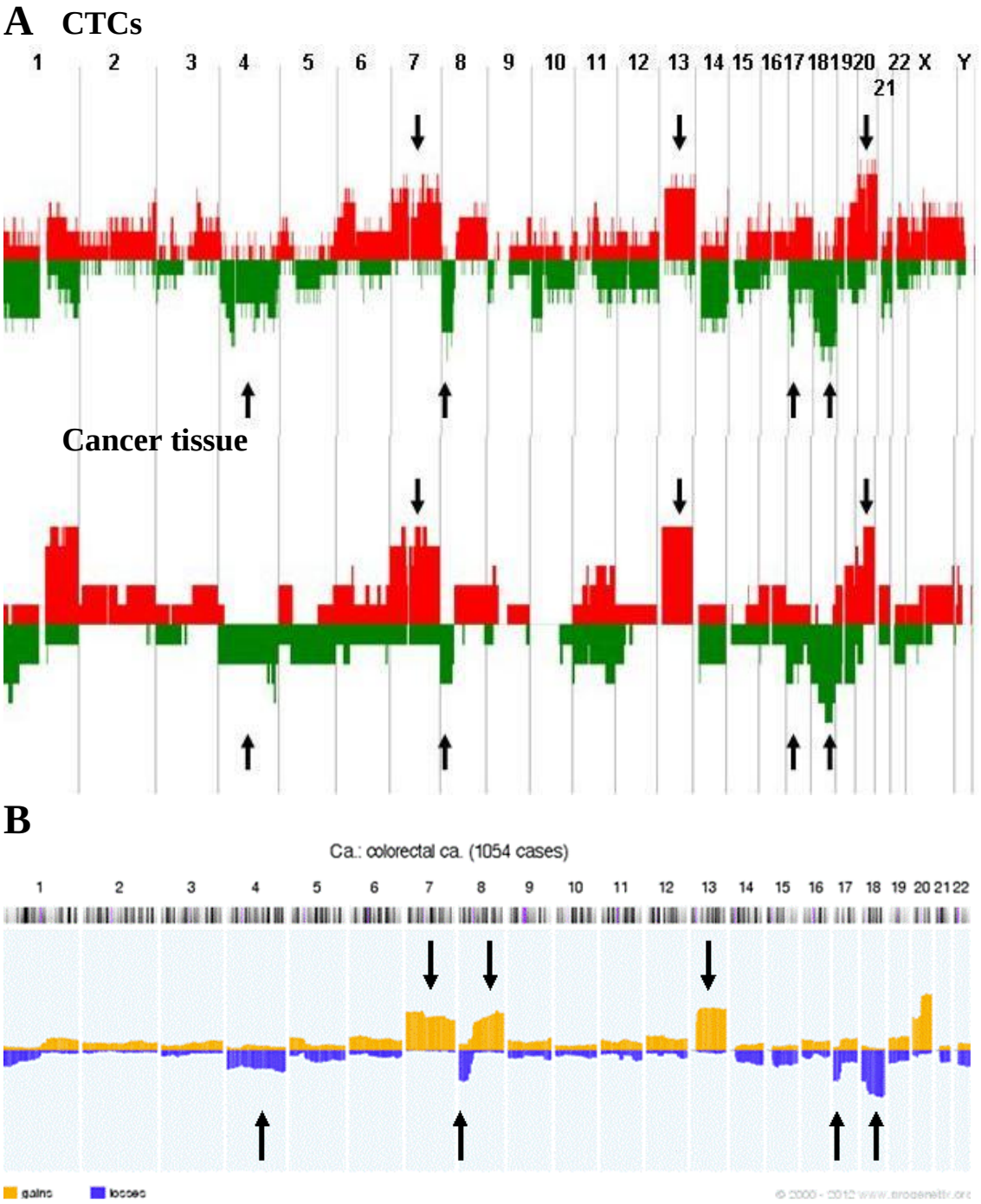


Figure 24: Accumulative penetrance plots of array-CGH data. A The penetrance Plots of pooled array-CGH results of CTC and cancer tissue samples are shown. Therefore, the human set of chromosomes (1-22, X, Y) was aligned and the frequency of detected aberrations was depicted. Arrows indicate typical CRC-associated chromosomal gains (red) and losses (green) found in both specimens. **B** For comparison of CGH profiles to database entries a penetrance plot provided by the university of Zurich (www.progenetix.org) is shown. The depicted plot includes CGH data of 1054 cases of CRC. Significant genomic aberrations shown in database were similarly detected in own analyses.

An orthotopic mouse model of cancer cell dissemination

As confirmed by own results, CTCs are rarely to find in blood samples of CRC patients. Therefore, we aimed to establish a model to mimic cancer cell dissemination. Animal experiments on cancer metastasis which inject tumor cells into the tail vein [119] or the portal vein[120] are little physiologic and omit initial steps of primary tumor growth and invasion. Ectopic animal models like subcutaneous tumor models [121] are easy to generate; however, most of them do not metastasize [122, 123]. Thus, we decided to establish an orthotopic mouse model of metastatic CRC because the microenvironment in the host organ is known to have an enabling effect on metastasis [124]. With the murine model it was intended to reproducibly obtain CTCs to a higher number than from human samples, because it is possible to take the total blood volume of a mouse through terminal heart puncture. Moreover, the disseminated cells would be of human origin; accordingly, possible contaminations with murine leukocytes would not influence the results of the target gene specific mRNA expression analyses.

With the help of a microinjection system human CRC cells were injected into the cecal wall of the animals. First trials were started with Balb/c nu/nu mice. The nude mice lack a thymus which makes them incapable to generate mature T-cells and makes them suited for xenotransplantation [125]. We decided to test different CRC cell lines including HCT116, LS174T, SW620, DLD1 and Colo205 for their potential of hematogenous dissemination in the nude mice. These cell lines were chosen because their use in orthotopic tumor models has already been described [126, 127]. More than 80 % of the animals developed orthotopic tumors. Additionally, lymph node metastases were frequently observed (Figure 25). No DTCs were found in the bone marrow. Solely implantation of HCT116 lead to macroscopic liver (26%; n = 19) or lung (5% n = 19) metastases (Figure 26).

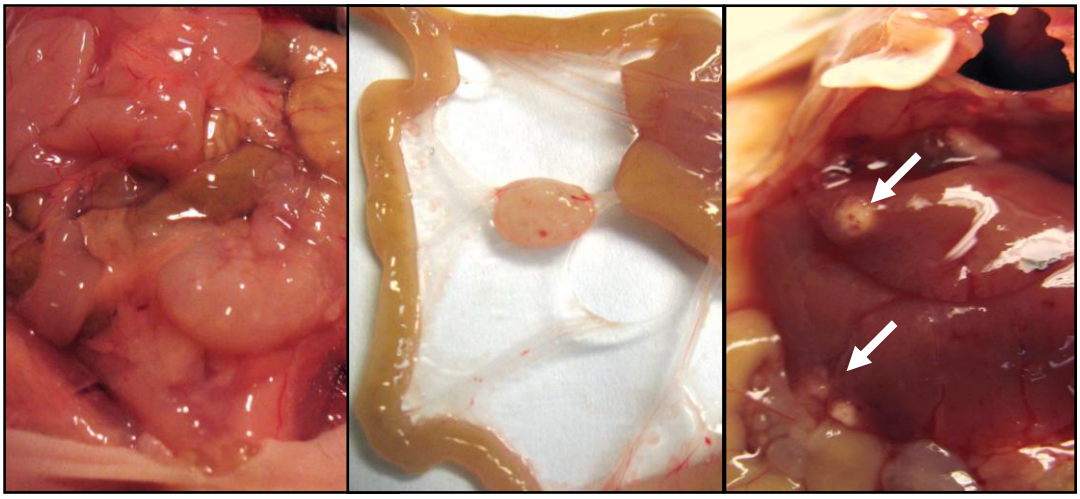


Figure 25: The orthotopic mouse model of CRC metastasis. The primary tumor (left), a lymph node metastasis (middle) and macroscopically visible liver metastases (right) of a xenotransplanted animal are displayed.

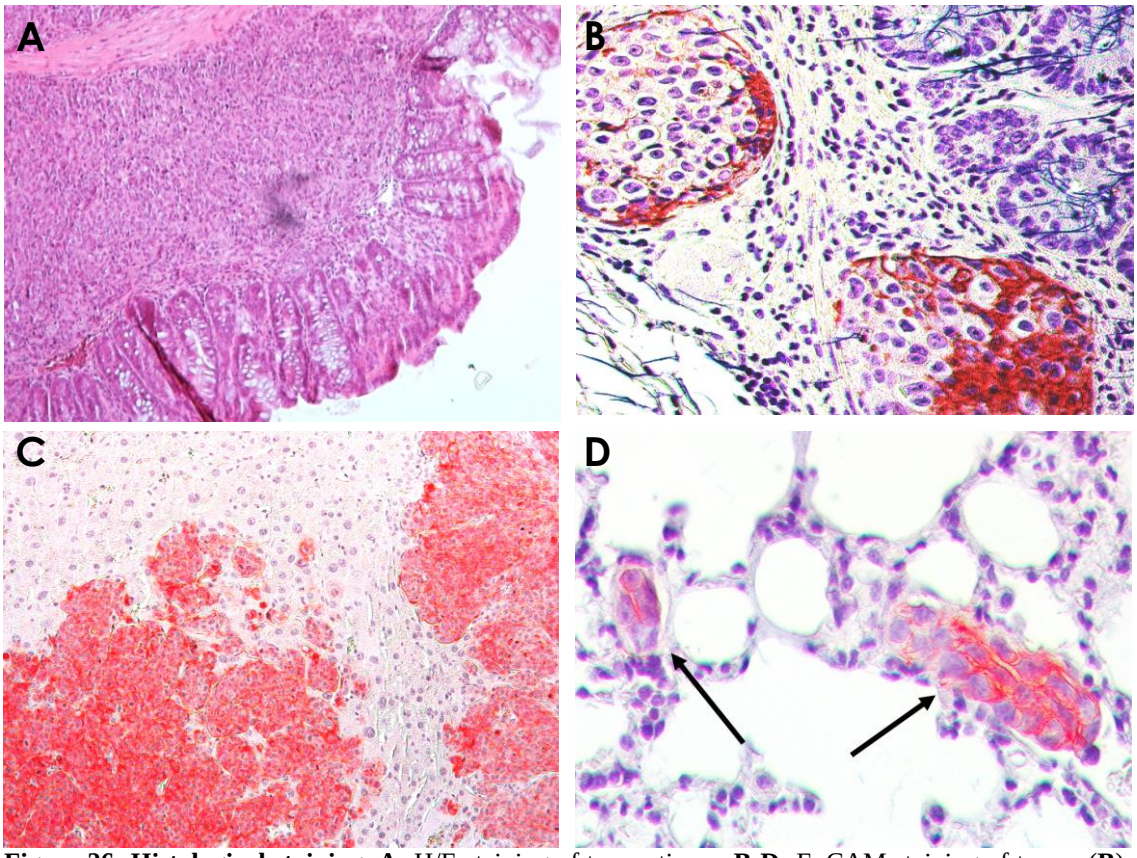


Figure 26: Histological staining. A: H/E staining of tumor tissue. **B-D:** EpCAM staining of tumor (**B**), liver metastases (**C**) and micro-metastases in the lungs (**D**) of a HCT116 xenotransplanted animal. The red color indicates tumor cell - associated expression of human EpCAM. Nuclei appear blue due to counterstaining with hematoxylin.

Of the tested cell lines, HCT116 seemed to be the only one with a tendency for hematogenous tumor cell dissemination. Therefore, further experiments were carried out only with HCT116. Immunohistochemistry and flow cytometry analysis of stained tissue samples from tumor, liver and lung metastases showed that the transplanted

HCT116 cell line retains EpCAM expression after xenotransplantation which makes EpCAM a suitable marker for the detection of CTCs and DTCs in blood and bone marrow (Figures 26 and 27).

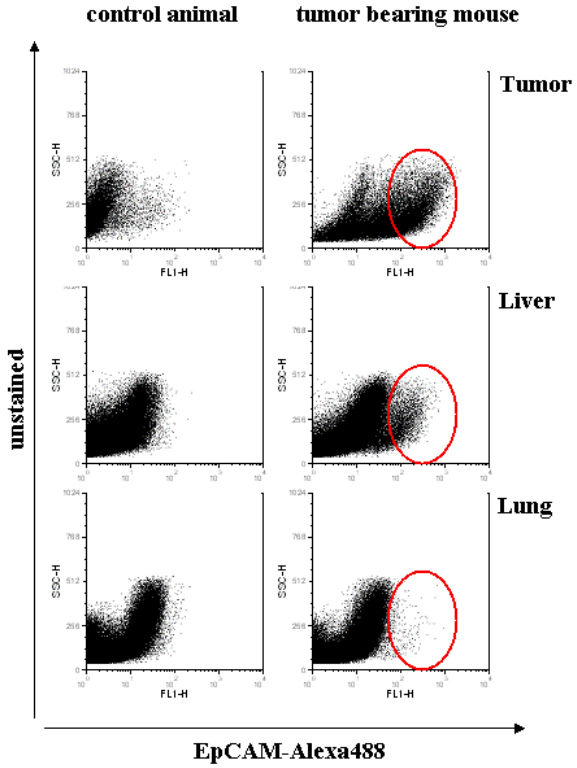


Figure 27: EpCAM expression of HCT116 after xenotransplantation. Single cell suspension of primary tumor, liver and lung of a xenotransplanted mouse and a mock transplanted control animal were stained with anti-EpCAM-Alexa-488 antibody and analyzed by flow cytometry. EpCAM-expressing tumor cells (as indicated) were found in all analyzed tissues of the xenotransplanted mouse.

29 animals that were xenotransplated with 2 x 10^6 HCT116 cells developed a primary tumor but CTCs were only observed in five mice. As we were surprised of such a low frequency, we supposed that the level of immunodeficiency in Balb/c nu/nu mice sufficed to clear CTCs out of the blood. Consequently, it was decided to change the mouse strain to NOD/SCID/gamma mice. Currently, these mice are among to the most immunodeficient strains available [128]. They lack mature T-cells, B-cells and natural killer cells and have defects in various cytokine signaling pathways. In addition, we decided to increase the potential for hematogenous dissemination by in vivo passaging of the cell line since this has been shown to augment the metastatic capability [129]. Thus, instead of original HCT116 we used a sub-cell line that was obtained from an HCT116-derived liver metastasis. This cell line is termed HCT116-P in the following.

To assess tumor development and the point in time of tumor cell dissemination HCT116-P tumor cells were orthotopically transplanted into NOD/SCID/gamma mice. Tumor growth and appearance of CTCs were monitored over time (Figure 28). All

animals developed primary tumors, liver metastases and lung metastases. Liver and lung metastases were assessed by histological staining (H/E staining). Liver metastases appeared between day 14 and 21, lung metastases developed between day 21 and day 28. Beyond that, the CTC amount in cardiac blood was determined as well. No CTCs were found before day 25 post tumor cell injection (ptci).

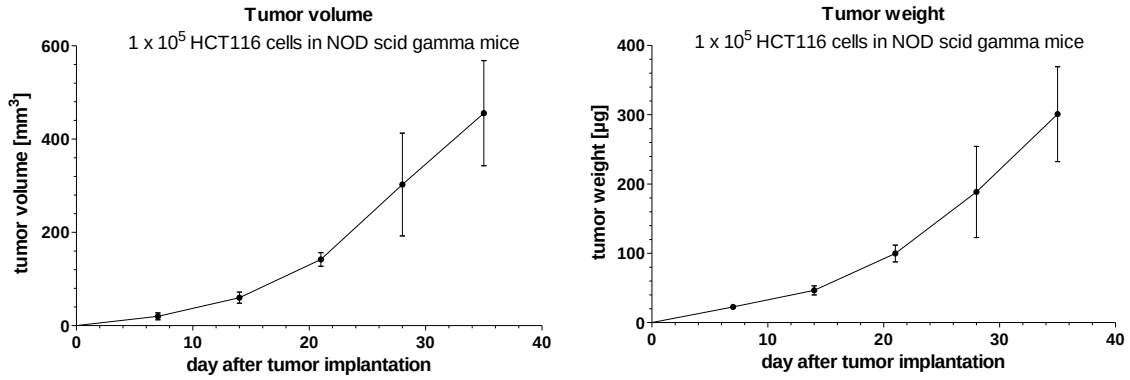


Figure 28: Tumor development. Graphs show the time-dependent increase of tumor size and weight.

Within 30-35 days all animals developed tumors of a mean weight of 0.3 g (Figure 28). At this stage lymph node metastases and macroscopic visible liver metastases were seen in all mice. A correlation of tumor weight to CTC amount was not found. No DTCs were found in the bone marrow.

It was possible to isolate CTCs from eight mice. The amount of CTCs varied between the animals from zero to several hundreds of CTCs. Isolated CTCs were viable and formed colonies when cultivated *in vitro* (Figure 29). Morphological characteristics and the expression of EpCAM confirmed tumor cell identity and human origin of the adhesive cells. Moreover, the CTC-derived cell line induced tumor growth in NOD/SCID/ gamma mice when injected subcutaneously.

Cancer progression can lead to the accumulation of fluid in the peritoneal cavity which is called acites. Xenotransplanted mice suffering from advanced cancer developed ascites containing thousands of tumor cells. Tumor cells that have disseminated into the ascites are termed AS-CTCs in the following.

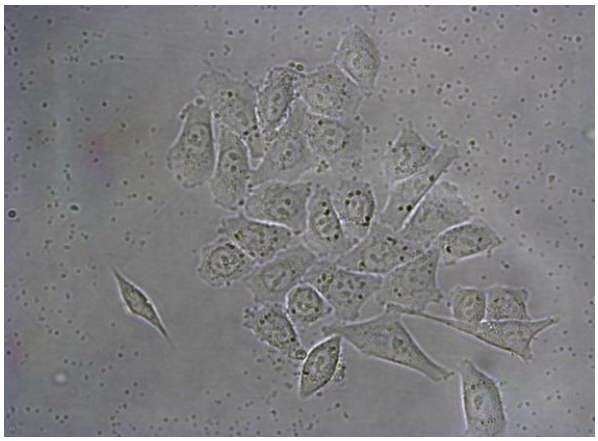


Figure 29: Colony formation of CTCs isolated from murine blood. CTCs isolated from mouse blood form colonies *in vitro*.

CTCs and AS-CTCs were stained6 with an EpCAM antibody for identification and isolation. Figure 30 shows that the staining intensity due to different EpCAM expression varies strongly among AS-CTCs. CTCs found in blood and ascites appeared often as cell clusters (Figures 30 and 31).

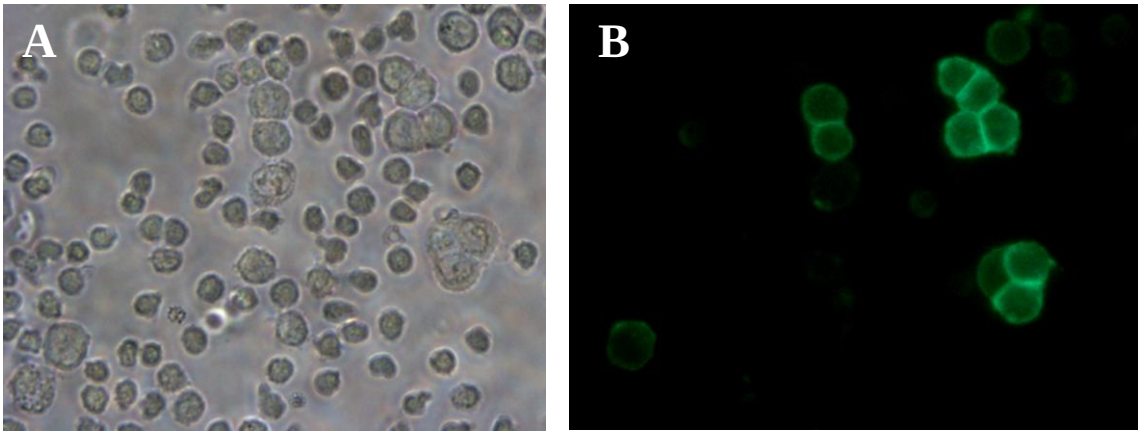


Figure 30: Fluorescence staining of AS-CTCs. AS-CTCs of a xenotransplanted animal were stained for membrane-bound EpCAM. **A** AS-CTCs and residual leukocytes in bright field. **B** Green fluorescence of anti-EpCAM-AlexaFluor 488 stained AS-CTCs. Different fluorescence intensity indicates variations in the protein expression levels among the AS-CTC.

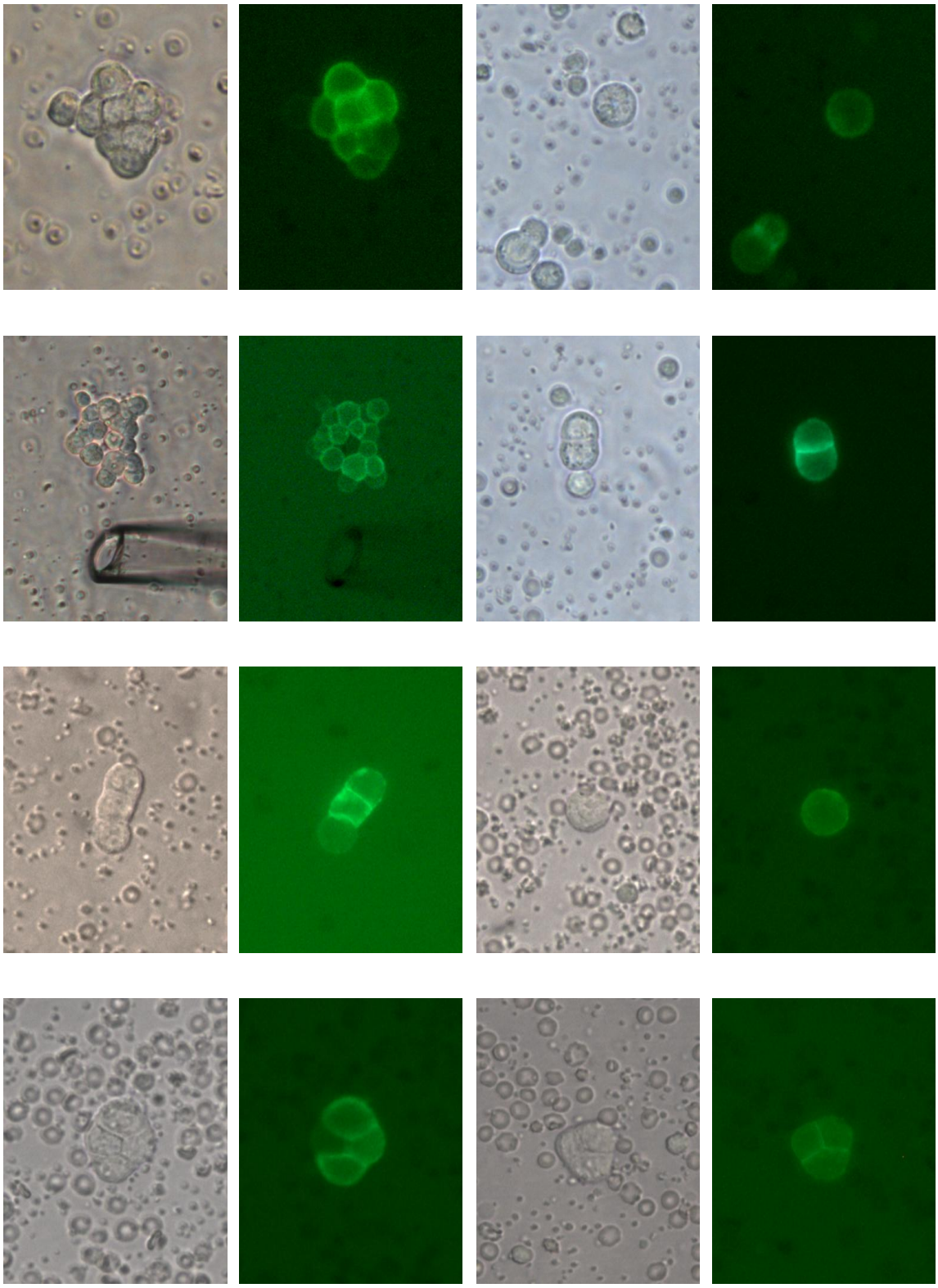
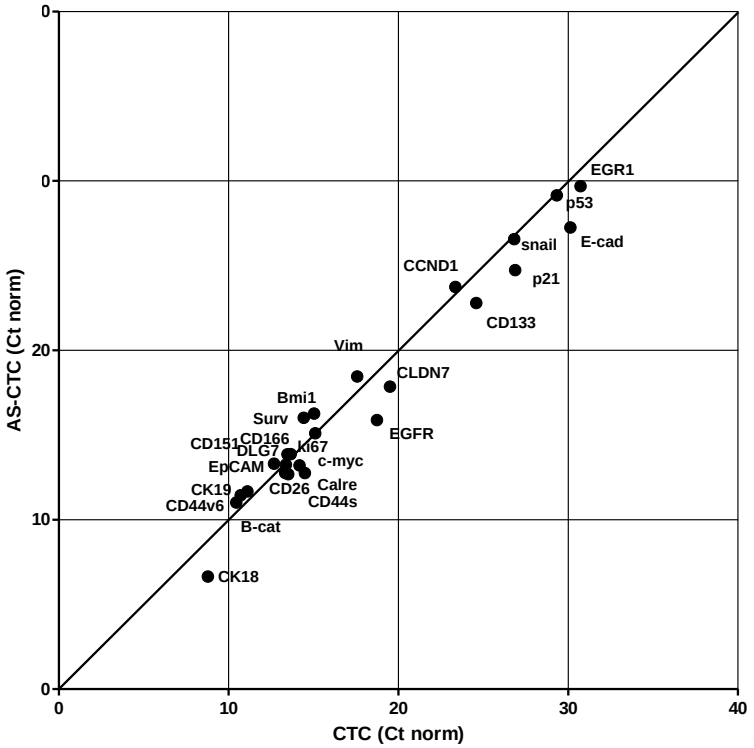


Figure 31: Fluorescence staining of CTCs. CTCs that were detected in murine blood are depicted in bright field and under green fluorescence after EpCAM specific fluorescence staining. CTCs detected in blood have a different appearance. Often different patterns of cell clusters were observed.

mRNA expression profiling reveals different gene expression patterns

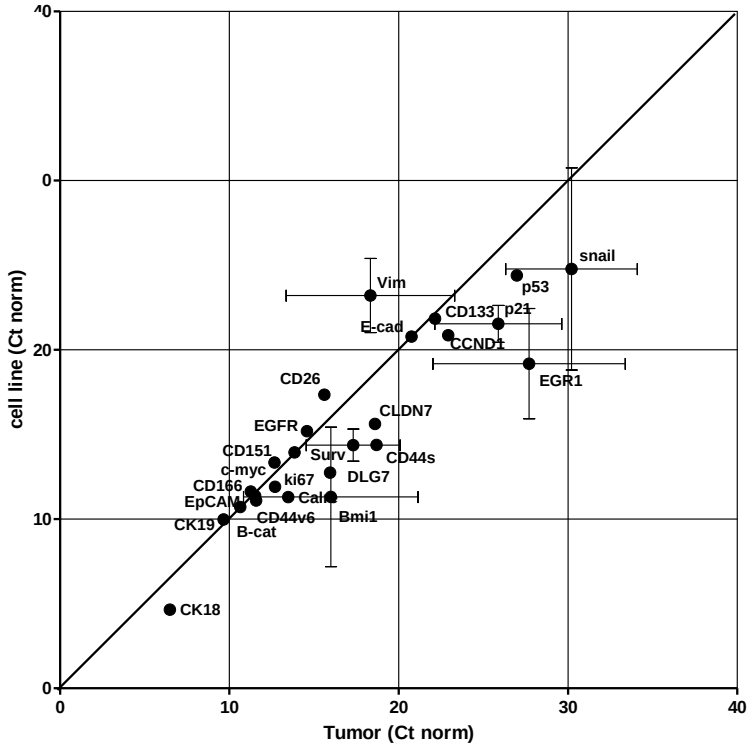
In each case about 15 cells of primary tumor tissue, of CTCs and of AS-CTCs were isolated and processed for mRNA expression analyses. After mRNA isolation and cDNA amplification the gene expression was evaluated with regard to cancer-associated biological traits like proliferation, cell cycle, stemness, metastasis and EMT. Reliable results of the conducted expression analyses were available for 17 tumors, 9 CTC and 18 AS-CTC samples obtained from eight mice. As expected, gene expression analyses of low cell numbers varied strongly. This might be attributed mainly to fluctuations in mRNA expression within individual cells. Therefore, common changes in gene expression require statistical evaluation and can only be detected if numerous samples are analyzed. Figure 32 gives an overview of the relative expression values. With regard to target gene expression CTCs resemble AS-CTCs. Comparing *in vitro* to *in vivo* culture of HCT116-P cells revealed most changes in the transcriptional profile. Genes that might be differentially regulated during tumor cell dissemination include E-cadherin, Snail, CD44s and EGFR as shown by comparing CTC with tumor mRNA expression.

A



R^2=0.97

B



R^2=0.81

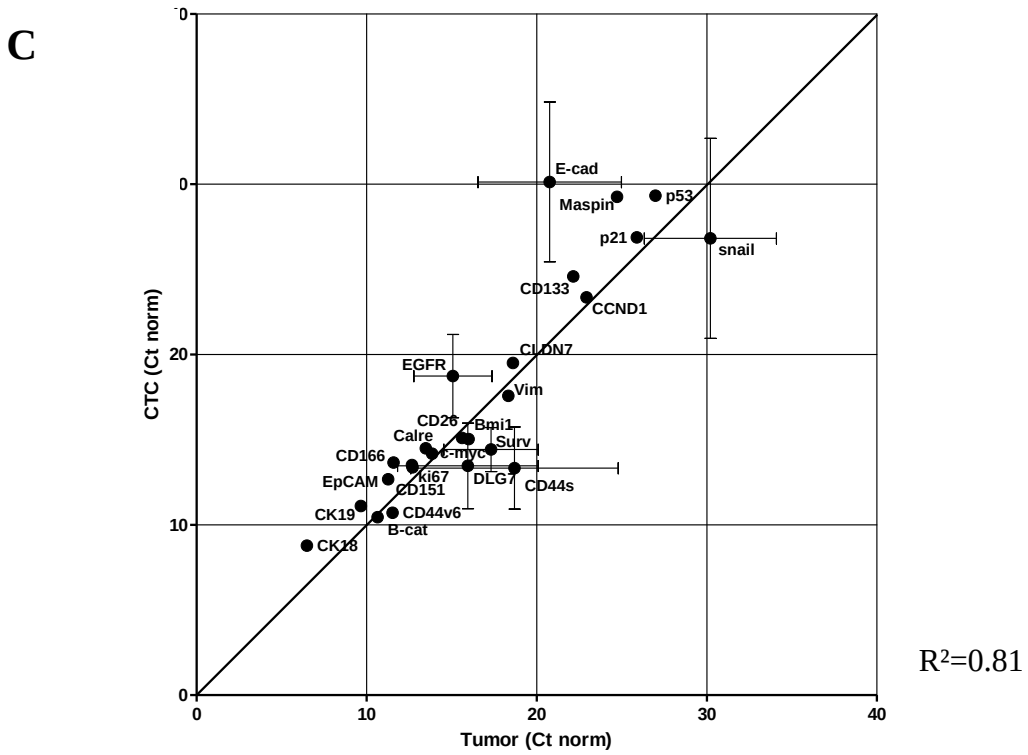


Figure 32: Comparison of mRNA expression. Transcript intensities are correlated between (A) CTCs and AS-CTCs, (B) unprocessed HCT116-P cells and tumor, and (C) CTCs and tumor. For reasons of clarity and comprehsibility standard deviation is only shown for selected genes. CTCs and AS-CTCs are transcriptionally similar. Unprocessed HCT116-P compared to HCT116-P-derived tumors display most differences in transcript intensity. High Ct values indicate a low gene expression.

CTCs maintain epithelial traits

To confirm the epithelial and human identity of the isolated CTCs expression of EpCAM, CK19 and CK18 was tested. Despite remarkable variations of transcript intensities the expression of the three assessed epithelial markers was constantly detectable. Figure 33 indicates that the expression of the epithelial markers is maintained during dissemination and no significant expression changes took place. This confirms that EpCAM, CK19 and CK18 are suitable and reliable marker genes for the detection of disseminated tumor cells.

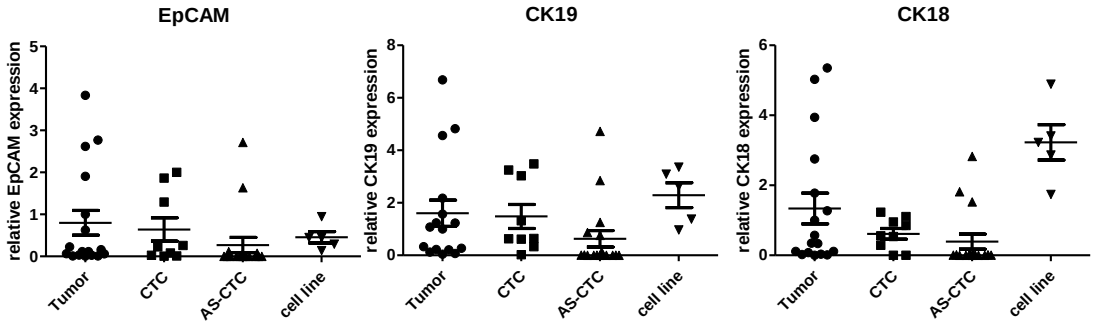


Figure 33: Expression of epithelial markers in CTCs. EpCAM, CK19 and CK18 expression of tumor tissue, CTCs and AS-CTCs normalized to β-actin are shown. No significant changes in mRNA expression among those genes were detected.

CTCs lack signs of EMT

To support the hypothesis that disseminating tumor cells undergo a change to a rather mesenchymal phenotype, specimens were tested for the expression of EMT-inducing transcription factors including Twist, Snail and ZEB1. Expression of Twist and ZEB1 was not detectable in any of the samples. Snail expression was found to be slightly up-regulated in CTCs (2 fold) and AS-CTCs (3 fold) samples; however, statistical evaluation did not confirm significance. Since EMT can be induced through numerous pathways, the expression of EMT target genes was analyzed, too. Although an average down-regulation of E-cadherin in CTCs (106 fold) and AS-CTCs (32 fold) compared to tumor samples was observed, data failed to reach statistical significance. N-cadherin and hFN expression were not detectable in the samples. Vimentin expression was detected in the specimens but expression was not significantly different among tumor and disseminated cells.

CTCs stop proliferation and up-regulate anti-apoptotic Survivin

To assess if disseminated cells persist in a dormant, non-proliferative state and if they are prone to undergo apoptosis the expression of Ki67, Cyclin D1, c-myc, β-Catenin, p53, p21 and Survivin was examined. Despite strong variations of mRNA expression levels, an up-regulation of Survivin was found in CTCs (4.6 fold, $p=0.004$) which was statistical significant (Student's t-test, $p < 0.05$) (Figure 34). The proliferation marker Ki67 was significantly down-regulated (10 fold, $p = 0.04$) in AS-CTCs, indicating a presumable proliferation arrest (Figure 34). No signs for differential gene expression were found for Cyclin D1, c-myc, β-Catenin, p53 or p21.

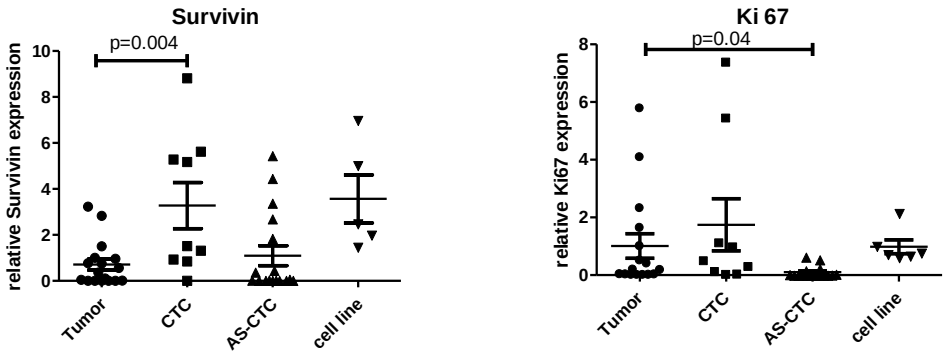


Figure 34: mRNA expression of Survivin and Ki67. Survivin and Ki67 expression was assessed in tumor, CTCs, AS-CTCs and in unprocessed HCT116-P cells. Survivin was found to be significant up-regulated in CTC samples potentially protecting CTCs from apoptosis. The down-regulation of the proliferation marker Ki67 in AS-CTCs seems to indicate a proliferation arrest in AS-CTCs as a consequence of long-term dissociation from the primary tumor site. Statistical significance was determined by Student's t-test. P - values below 0.05 were considered as statistical significant.

No significant changes in the expression of stem cell markers in CTCs

Samples have been analyzed for the expression of stem cell-associated markers. Expression of Olfm4, Ascl2 and Lgr5 has not been detected in the specimens. In case of CD133, CD166, Bmi1, EGR1 and DLG7 no significant differences in gene expression were found.

Most metastasis-associated genes are equally expressed in CTCs and tumor samples

The expression of EpCAM, CD44v6, CO029, CD151 and CLDN7 facilitates metastasis formation in CRC [130]. Therefore, we aimed to evaluate the expression of these genes in our mouse model. However, no significant differences in the expression of EpCAM, CD44v6, CD151 and CLDN7 were found. The expression of the tetraspanin CO029 was not detected. Interestingly, CD44s was found to be significantly up-regulated in CTCs (p= 0.017; 9.4 fold) and in AS-CTCs (p=0.03; 8.4 fold) as compared to tumor samples (Figure 35). Additional metastasis-associated genes including MMP7, CXCR4 and CD26 seemed not to be differentially regulated. EGFR was found to be down-regulated (5 fold) in CTCs compared to tumor samples but failed to reach statistical significance (Figure 35).

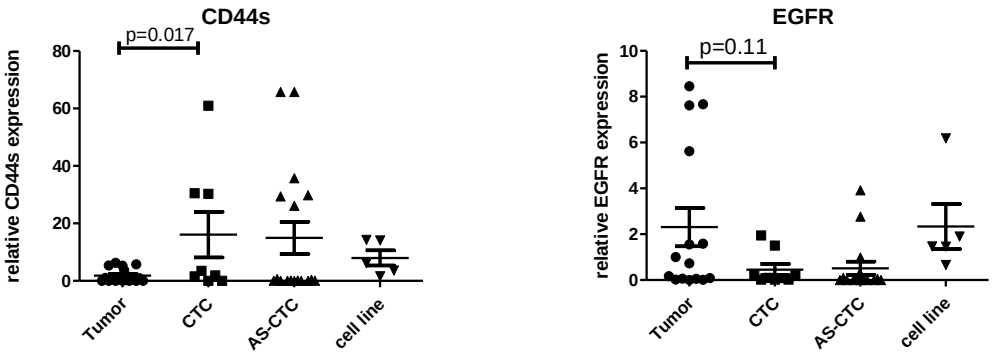


Figure 35: mRNA expression of CD44s and EGFR Relative expression of CD44s and EGFR was measured by RT-PCR. Statistical significance was determined by Student's t-test. P - values below 0.05 were considered to be statistical significant. Significant differences of CD44s mRNA expression were found between tumor and CTCs specimens. EGFR expression failed to reach statistical significance among the different subgroups.

In summary, our mouse model of metastatic CRC provides the opportunity to study differential gene expression during the process of tumor cell dissemination. We found that the expression of epithelial markers including CK18, CK19 and EpCAM is not significantly altered after the dissemination. Although down-regulation of E-cadherin seems to support metastasis, convincing signs of EMT in the disseminated cells were not found. However, the detection of differentially expressed genes like Survivin and CD44s, that were found to be up-regulated in CTCs, might allow first conclusions about the mechanism of tumor cell dissemination.

Discussion

This thesis provides comprehensive data about the detection, isolation and characterization of CRC - derived CTCs and DTCs. Beside an own study on CTC incidence in blood samples of CRC patients, we provide intriguing data about genomic and transcriptional profiles of CTCs and DTCs. In addition, we established an orthotopic mouse model to mimic cancer cell dissemination. To my knowledge a comparable wide-ranging approach to elucidate the molecular mechanisms of tumor cell dissemination has never been reported before.

Compartmental differences of CTCs in CRC

To study the CTC amount in our patient cohort we used the FDA-cleared CellSearch™ system. We analyzed blood samples from the CVBC, which were either obtained from the *internal jugular* or *subclavian* vein, and blood samples from the MVBC, which were collected from a tumor-draining mesenterial vein. Blood was taken shortly before surgical tumor resection and allowed the evaluation of compartmental differences of CTC distribution in both venous compartments. We could confirm that the detection of CTCs in central venous blood correlates to the stage of disease [53]. CTCs were found rarely and in low numbers in blood samples of patients with early stages of CRC. Therefore, the use of CTC detection systems for prognostic and therapeutic purposes might be largely limited to advanced stages of CRC. The low number of CTCs that are detected in early cancer stages is usually below the dynamic range required for measuring treatment response.

Following the pathways of tumor cell dissemination in CRC we have compared the distribution of CTCs in central and mesenteric venous blood. CTCs were found to be more frequently present and with higher numbers in MVBC than in CVBC. Previous studies [131-134] that have analyzed the presence of CTCs in the mesenteric venous blood of CRC patients are in line with our findings. However, those studies relied mainly on RT-PCR to detect the expression of CEA or cytokeratins and did not allow the exact enumeration of CTCs. Our data supports the cascade theory in CRC assuming a stepwise tumor progression with the liver acting as a filter for CTCs [50]. The detection of CTCs in central venous blood indicates that the filtering function of the liver seems

to be incomplete. We found an increased quantity of CTCs in CVBC in patients with CTCs in the MVBC suggesting an amount-dependent filtering of CTCs in the liver. The development of overt metastases causes further dissemination of CTCs into the circulation which is supported by the findings that the CTC amount in CVBC increases considerably in patients with stage IV disease and that the presence of liver metastases was found to be a strong and independent factor associated with CTC detection in the CVBC but not in the MVBC. Although various studies [53, 132] regarding the association of clinicopathologic variables with the detection and count of CTCs in mesenteric and central venous blood already exist, there has been no concomitant evaluation of CTC incidence in CVBC and MVBC using the standardized CellSearch™ system. An elevated preoperative CA 19-9 level was found to be independently correlated with the presence of CTCs in the CVBC. CA 19-9, which has been widely used as a tumor marker, is a sialylated Lewis blood group antigen and a ligand for the endothelial cell adhesion molecule E-selectin [135]. Several studies have shown that the preoperative serum CA 19-9 levels are associated with a poor prognosis for patients with CRC [136, 137]. Beside the function of CA 19-9 in adhesion of CTCs to the endothelium, our data might suggest that CA 19-9 could possibly be involved in the detachment of tumor cells from the primary tumor and in CTC persistence in the systemic circulation.

The prognostic relevance of the detection of CTCs in the mesenteric compartment is still controversial. A meta-analysis [84] of the few available studies was not able to confirm a significant association of mesenteric CTCs with disease recurrence and survival whereas CTC detection in the peripheral/central blood was a strong predictor of poor outcome. Our data revealed that CTC detection in the MVBC is only weakly correlated to the assessed clinicopathological variables and is in line with the inconsistent finding of previous studies. Therefore, we suppose that most of the CTCs that are found in the MVBC are not of clinical relevance. Probably, most of CTCs that are shed into the circulation can not adapt to the foreign microenvironment and die quickly [41-43]. However, CTCs that are found in the CVBC have proved their ability to survive in the systemic circulation and have accomplished a further decisive step on the way to malignant progression.

Establishment of a reliable CTC enrichment and detection method

mRNA expression studies on CTCs and DTCs are essential for understanding the biology of the initial steps of the metastatic cascade. Since standardized methods for CTC/DTC isolation are lacking, own approaches for single cell analyses had to be developed and evaluated. The right choice of suitable CTC markers appeared to be a major problem. Since intracellular proteins such as cytokeratins require the permeabilization of cells for staining, they were not eligible to obtain intact cells required for mRNA expression studies. We decided to focus on EpCAM as epithelial surface marker. EpCAM is widely used for the detection of CTCs in blood samples of cancer patients since it is absent on leukocytes. Also the CellSearch™, the only FDA-approved, standardized system for CTC enumeration in various cancers makes use of EpCAM as initial CTC marker [85]. Nevertheless, EpCAM remains controversial because it is described as a cell surface protein with functions in cell-cell adhesion that might counteract tumor cell shedding. In 2009 Sieuwerts et al. [138] reported that some breast cancer cell lines even lack EpCAM expression. In addition, some authors [40, 139] claim that epithelial genes such as cytokeratins and EpCAM might be down-regulated in the course of EMT which is assumed to be critically linked to tumor cell dissemination. Although a general presence of EpCAM on CTCs might be uncertain, a total loss of EpCAM on CRC - associated CTCs has not been shown so far. Despite concerns EpCAM is presently at least to our opinion the best suited marker to distinguish epithelial cells from leukocytes in human blood samples. Therefore, we have decided to use EpCAM as an eligible marker for CTC detection.

Comparing the efficiency of our manual CTC enrichment approach to the standardized CellSearch™ system revealed that CTCs were detected to a lower number and less frequently applying the manual protocol. Whereas the CellSearch™ applies ferrofluidic nanoparticles and strong electromagnets to enrich CTCs, our method depends on leukocyte depletion by relative large magnetic beads that are removed with simple ferromagnets. In addition, our approach requires searching for CTCs under a fluorescence microscope by the unaided eye. This might not be as efficient as it can be achieved with the sensitive camera system that is used in the CellSearch™. Nevertheless, our manual CTC enrichment method enabled the detection of intact CTCs and with the help of the micromanipulator it has been possible to directly isolate single CTCs without a strong leukocyte carryover. Establishing a functional CTC

detection and isolation protocol has been decisive for the following study on mRNA expression in CTC samples.

A mRNA expression study on CRC - associated CTCs and DTCs

At least to our knowledge, here, we describe for the first time a study on the mRNA expression of almost 50 genes in nearly pure fractions of CRC-associated CTCs and DTCs. CTCs/DTCs are rare and only to find in a strong background of leukocytes or bone marrow cells, which is a major obstacle for their molecular analysis. The impurity of the CTC/DTC fractions used in previous reports on their transcriptional signatures has been a substantial drawback of these studies [100-102]. Analyzing rare CTCs in a strong background of several thousands of leukocytes is mainly limited to the detection of genes that are exclusively expressed by CTCs and can therefore only be used to define novel CTC markers. However, an unbiased evaluation of mRNA expression in CTCs/DTCs requires pure cell samples. With our micromanipulator-based method that allows the pipetting of single cells, we aimed to obtain CTC/DTC samples free of any leukocyte carryover. Despite promising results with blood samples spiked with cancer cells, it was, however, not possible to isolate CTCs/DTCs from patients' blood samples lacking any CD45 mRNA signal. CD45 was actually intended to measure leukocyte carryover; however, some researchers recently reported co-expression of CD45 with epithelial markers in CTCs of some patients with different kinds of cancer [140]. At the moment we can only speculate about a putative contribution of CTCs to the detected CD45 signal. Further efforts are needed to clarify if CD45 is attributed exclusively to leukocytes or if it is at least occasionally expressed on CTCs as well. In addition, it was recently reported that CTCs might avoid NK cell-mediated lysis through the aggregation with platelets [47, 48]. Moreover, Pawelek et al. have suggested that CTCs might fuse with BMDC in the blood [49]. Both scenarios could lead to the CD45 detection in the specimens. Probably, our CTC/DTC samples are not absolutely free of any leukocyte carryover; nevertheless, the leukocyte background has been considerably reduced compared to other studies [100-102]. The detected CTC/DTC expression profiles are distinct from profiles of leukocytes or cells of the bone marrow and show a unique expression signature. Epithelial markers such as EpCAM, CEA, Calreticulin, CK18 and CK19 indicate a substantial amount of cancer cells in the evaluated samples. Generally,

a strong heterogeneity among the CTC/DTC specimens was observed which might indicate the existence of various CTC/DTC subpopulations.

Although mRNA expression analysis of single cells would be feasible, we decided to pool CTC populations to obtain more valid data and to increase the chances to detect even lowly expressed genes. Studies on gene expression in single cells or small cell pools are challenging until today. A large scattering of detected mRNA expression levels has to be expected since transcription is rather a stochastic than a continuous process [141]. The inherent randomness in transcription impedes the precise phenotypic determination of individual cells. Particularly, the noisy transcription of lowly expressed genes requires large numbers of samples and statistical evaluation to state significant trends.

CTC samples

In this pilot study CTC samples from only ten patients were evaluated, common findings are therefore limited but some remarkable aspects are worth discussing. Interestingly, some genes, including CD47, CD44s and Vimentin seem to be up-regulated in CTCs when compared to tumor tissue, however rather similar expressed when referred to metastasis samples. It might be conceivable that the expression of some metastasis-associated genes in CTCs resembles the metastases rather than the primary tumors. For instance high expression of CD47, which might be involved in tumor immune escape [105], might be necessary for CTC survival in the circulation but could be favorable for metastatic growth as well. Probably, metastases maintain at least to some extend the expression profiles they already gained during the early steps of tumor cell dissemination. E-cadherin is another example supporting this hypothesis. E-cadherin was found to be down-regulated in the CTC samples when compared to primary tumors, however similarly low expressed as detected in metastases samples. Lower expression of E-cadherin in metastases compared to matched primary tumors was already reported for CRC [142]; hence, the reduced expression of cell-cell adhesion molecules seems to be crucial for tumor cell detachment. However, our results do not confirm a significant contribution of EMT to tumor cell dissemination. It is widely reported that EMT might be a prerequisite for metastasis [39] but our CTC samples seem to lack expression of various putative inducers of a mesenchymal transition including Twist and Snail. Furthermore, an up-regulation of mesenchymal genes was also hardly detectable. One might argue that our CTC identification method relies exclusively on

the epithelial marker EpCAM and therefore could miss CTCs that have undergone EMT and lack expression of epithelial proteins. Recent studies [139] with the ISET system, a CTC detection method depending on size-dependent filtration, have shown the presence of non-small-cell lung cancer-associated CTCs that did not express epithelial markers. However, at least to my knowledge neither the existence of CRC-associated CTCs lacking the expression of epithelial markers nor the contribution of such cells to cancer staging or prognosis has been reported so far. In addition, if mesenchymal-transformed tumor cells are able to initiate metastasis is highly controversial at all [57 143]. On the other hand the impact of $EpCAM^+$ - CTCs on cancer prognosis was confirmed several times [86 87].

It was reported, that CTCs are frequently apoptotic [144]. We did not detect the differential regulation of apoptosis related genes such as Bax, Survivin, p53 and p21 in our samples. However, our CTC samples were obtained from tumor or metastasis draining vessels and not from peripheral blood samples; therefore most of the analyzed CTCs have probably just detached from the solid cancer tissue and represent only a snapshot of an assumed heterogenic CTC population. Apoptosis might occur delayed and after a longer exposure to shear stress and the loss of cell-cell contacts. Interestingly, we did not find an up-regulation of metastasis-associated genes, including CD44v6, CO029 and CLDN7 in the CTC samples. Despite convincing reports [130, 145] that these genes indicate aggressive tumor growth and metastatic spread, their expression seemed not to be attributed to the analyzed CTC samples.

We found the proliferation marker Ki-67 absent or down-regulated on CTC samples which is in accordance to a study about breast cancer-associated CTCs [146]. Although we can not exclude that proliferating CTCs exist, a proliferative arrest might be useful for CTCs to adapt to the foreign microenvironment in the blood. Interestingly, we did not detect the down-regulation of β-catenin which would additionally indicate a low proliferation of CTCs since it is a main part of the Wnt signaling pathway [147]. However, regulation of the Wnt pathway might be achieved post-transcriptionally as well [148].

DTC samples

It was only possible to include two samples of DTCs isolated from the bone marrow of CRC patients in this pilot study. We were able to identify DTCs due to EpCAM expression amongst a strong excess of bone marrow cells. Detection of CK18 confirmed the epithelial origin of the DTCs. Despite CK20 is the most commonly used

marker for DTC detection in RT-PCR analyses [90], no CK20 mRNA expression was detected in our DTC specimens. Together with the lost CK19 expression this might indicate signs of an occurred de-differentiation in the cells and questions the use of CK19 and CK20 as DTC markers. The gene expression profile of DTCs seems to be distinct from their corresponding primary tumors or metastasis. In particular with regard to systemic therapeutic approaches the molecular characterization of DTCs might have enormous therapeutic impact. For instance, DTCs, that we have analyzed, seem to lack EGFR expression which might mediate chemotherapy resistance and explain limitations of anti-EGFR based therapies in systemic treatment.

The proliferation marker Ki67 was also not detected in DTCs. This is in line with other reports that describe DTCs mostly as non-proliferative and dormant cells [146, 149]. However, DTCs might be able to escape proliferative dormancy when supplied with appropriate growth factors as shown by cell culture experiments [150]. DTCs that are hidden throughout the whole body are highly suspicious to cause cancer relapse even years after primary tumor resection [58]. Thus, to reduce cancer reoccurrence therapeutic regimens are required that effectively target DTCs. Since the molecular traits of the primary tumor are probably distinct from DTCs, the development of successful therapies calls for further ambitious studies on the characterization of DTCs.

Genomic characterization of CTCs

The present study combined the standardized CTC-enumeration of the CellSearch™ system with the concomitant genotyping of clearly defined and detected CTCs. Single CK-positive stained cells were isolated with a micromanipulator mounted on a fluorescence microscope for the subsequent global amplification of the gDNA. To evaluate the possibilities to characterize CTCs on the genomic level we searched for point mutations in *Kras*, *BRAF* and *TP53* as well as for signs of MSI. In addition, some cells were used for global array-CGH analysis. Point mutations, MSI and chromosomal instability have been associated with survival, prognosis or therapy response in CRC. For instance, mutations of *Kras* and *Braf* indicate a lack of therapy response to anti-EGFR monoclonal antibodies like cetuximab and panitumumab [151-153] which is caused by the constitutive activation of the Ras-MAPK pathway [154]. Mutated *TP53* has been associated with survival [155, 156] and therapy response [157]. MSI and specific chromosomal

losses or gains were found to be of prognostic relevance [158-162] as well. Therefore, the genomic characterization of CTCs might provide an opportunity to improve prediction of prognosis and to adapt therapy regimes to the genetic aberrations of individual tumors. With the present study we were able to show that epithelial cells detected with the CellSearch™ carry CRC-associated mutations and we provide evidence that these cells are of malignant origin. Despite a limited number of samples we detected remarkable inter-CTC heterogeneity in CTC fractions obtained from single patients. Of the 17 patients from whom we were able to analyze more than one CTC a genetic disparity among CTCs of a single patient was found in six cases (37.5%). Genomic heterogeneity of CTCs seems to be a common phenomenon and an evidence for genomic instability and clonal diversity in CRC. A study from Fehm and colleagues [163], that mainly included cases of breast cancer, reported heterogeneous CTC pools in 13 of 20 cases (65%). However, they used *in-situ* hybridization for the detection of aneusomie. Since their approach covers more potentially mutational sites, the higher detection rate might thus be explained. The presence of genomic heterogenic CTC populations further complicates the selection of appropriate therapy regimens and might require the analysis of a number of representative CTCs when individual tailored therapy is intended.

In addition, we aimed to evaluate to what extend the detected genomic aberrations correlate among CTCs and matched cancer tissue. In twelve of 26 cases we found disparate genotypes for CTCs compared to cancer tissue. Intriguingly, a high portion of disparate genotypes (CTCs vs. cancer tissue) was found for *Kras* and *BRAF* whereas disparity was less frequently observed for *TP53*. Typically, inactivating mutations of tumor suppressor genes like *TP53* are found homozygous [164]. Oncogenes like *Kras* or *BRAF*, however, often carry activating mutations in only a single allele and are therefore heterozygous [165]. One might argue that the applied single cell gDNA amplification technique can not reliable detect heterozygous mutations. Due to the low amount of gDNA of a single cell some DNA fragments could fail to be amplified and thus cannot be detected in sequence specific PCR. Therefore, sequencing results might be misleading in heterozygous situations if only one of two different alleles can be analyzed. Generally, it can not be excluded that allelic losses occur during the gDNA amplification. Therefore, it is inevitable necessary to define the rate of amplification-associated allelic losses. Since no own experiments have been performed on this, we refer to the work of Schardt, et al. [166] who reported a rate of allelic losses due to

technical issues of 9.8 % in their control. Although the processing of blood samples with the CellSearch™ system differs from the methods that Schardt and colleagues have used, we expect a comparable rate of allelic losses because the same amplification protocol has been applied. The global amplification of gDNA may have caused allelic losses that left heterozygous *Kras* or *BRAF* mutations in CTCs unrecognized. However, the disparity between the detected mutations in CTCs and corresponding tissue cannot be explained exclusively by technical reasons. We believe that the disparity probably results from the use of marco-dissections to analyze the malignant tissue. Despite a presumptive monoclonal origin tumors are heterogeneous and consist of various subpopulations [106]. Minor subclones that harbor only the wild-type alleles might remain unrecognized when thousands of cells were pooled for analysis. In addition, it is conceivable that specific genetic changes might have supported tumor cell dissemination. For instance, the loss of genes involved in adherence or the amplification of genes related to invasiveness would probably favor the dissemination process. To our knowledge there are no comparable studies about genomic heterogeneity in CRC associated-CTCs. Most studies analyze genetic disparity between primary tumors and various sites of metastases. Findings indicate, despite considerable inter-study variations, disparate point mutations for *Kras* in 0-60% [98, 167-169 170] and for *TP53* in 20-60% [171, 172] of cases. There are a few studies describing genetics of DTCs in bone marrow [97, 173-175]. The frequently observed genomic disparity in DTCs in particular in early stages of cancer, they observed, is often explained with postulated early tumor cell dissemination [99]. Tumor cell diversity is still large in early stages since genomic instability will lead to the development of various subclones. Once the tumor reaches full malignancy the most aggressive clone will prevail. An early dissemination and the subsequent parallel progression of DTCs, which further contributes to the divergent development, might explain the detected genomic disparity. However, it is believed that CTCs can not survive for long time in the circulation. Shear stress [43], anoikis [44] and the easy access of the immune system [176] are thought to be the main reasons for the short persistence of CTCs in the blood. The high number of genetic disparity between CTCs and tumor tissue, that we have detected, is fairly surprising and in contrast to a study from Khan et al. [177]. They studied *TP53* mutations in primary tumors, liver metastases and CTCs enriched by magnetic beads out of blood samples from CRC patients. They reported that if a *TP53* mutation was detected it was found invariably in all tumor and liver metastasis samples from a single patient and in eight of

19 cases it was found additionally in the CTC sample. However, their method could not determine if *TP53* wild-type tumor cells were also present in the enriched CTC fraction or in *TP53* mutated carcinomas since they approach did not allow single cell analysis. Our CTC isolation protocol provides the opportunity to analyze single CTCs without any leukocyte carryover and is, thus, well-suited for CTC- genotyping. CTCs can be obtained repeatedly and relatively noninvasive from blood samples, therefore, information about their genomic profile might contribute to treatment decision and the estimation of prognosis. For instance, information about the mutational status of *Kras* might be crucial for individual therapeutic intervention. The failure of anti-EGFR based therapies was reported for colorectal tumors harboring mutant *Kras* [152] or *Braf* [153]. We were able to show that the mutational status of *Kras* and *BRAF* can differ among CTCs and matched cancer tissue. Cancer cells might have disseminated at an early point in time and have developed in parallel to the tumors they have derived from. Therefore, we conclude that genomic analysis of single cancer tissue specimens will not be sufficient to state about the genomic profile of systemic cancer. Moreover, it is even conceivable that genotypic changes of CTCs might occur during cancer progression. This was shown by Meng and colleagues [178]. They observed in nine of 24 breast cancer patients an acquired Her-2 gene amplification in CTCs although the primary tumor was initially Her-2 negative. Conversely, the emerge of Her-2 negative CTCs in Her-2 positive breast cancers subjected to anti-Her-2 therapy was reported by Hayes and collegues [179].

In conclusion, here, we describe that the comprehensive genomic characterization of single CRC-associated CTCs is feasible. Cancer tissue biopsies consist of an average of malignant and non-malignant cells of a tumor and usually neglect individual cancer subclones. However, since tumors might consist of heterogeneous cancer cells, the single cell analysis could be favorable. We showed that point mutations, MSI and chromosomal aberrations that might be important for prediction of survival and therapy response can be evaluated in single cells. In addition, the identification of genotypes that could be associated with tumor cell dissemination and ectopic survival might open new opportunities to develop effective treatments.

An orthotopic mouse model of cancer cell dissemination

CTCs are extremely rare and difficult to detect in blood samples of CRC patients. To mimic metastatic spread, we have established an orthotopic mouse model of cancer cell dissemination in NOD/SCID/gamma mice. Our CRC model comprises all relevant steps of metastasis including lymphogenic and hematogenic tumor cell dissemination and the successful formation of metastases in liver and lungs. It was possible to isolate CTCs from cardiac blood and tumor cells from the ascites of some animals. CTCs were viable and could be cultured and expanded *in vitro*. Moreover, subcutaneous injection of CTC-derived cells confirmed their tumorigenicity.

Clearly lesser CTCs were detected in the cardiac blood in our CRC model than reported for othotopic xenograft models of breast [180]and prostate [181] cancer. Most murine models determine CTC numbers in blood obtained from heart puncture. In case of an orthotopic colon cancer model the venous blood that leaves the tumor needs to pass the liver before it reaches the heart [50]. However, most of the draining blood of prostate or mamma carcinomas is drained via the *vena cava* without having to pass another organ which may act as a filter for CTCs. The detection of lower CTC numbers in our CRC model confirms the strong filter effect of the liver that we have already postulated in the human setting. Moreover, metastases seem to significantly contribute to overall cancer cell dissemination since no CTCs were detected before liver metastases have formed. We did not find any correlation of the tumor size to the CTC number in the xenograft model. Although this is in line with previous findings [180, 181], the presented cancer model does not allow for the accurate determination of tumor burden throughout the animal. Micrometastases might have established at unknown sites distributed all over the body. Applying constantly marked tumor cell lines would be desirable for further studies. Cancer cell lines with intrinsic GFP or Luciferase expression might facilitate an exact tumor and metastasis quantification and might enable an unbiased CTC count that does not depend on the expression of epithelial genes.

Most of the studied genes were found to be similarly expressed in CTCs and solid tumor samples. This is in accordance with a previous report about CTCs in an othotopic prostate cancer model [181]. We share the opinion of Helzer and Barnes that the biological features necessary for metastatic spread might be already inherent in the aggressive cell lines used in both models. This might explain the limited number of differentially regulated genes that were detected. Interestingly, Howard and

colleagues [182] established a CTC cell line from the same orthotopic PC-3 prostate cancer model. They report a down-regulation of cell adhesion molecules including E-cadherin in the CTC-derived cell line. Although, not statistically significant, a clear trend for the down-regulation of E-cadherin in CTCs was also observed in our study. Together with similar findings in human CTC samples E-cadherin down-regulation might be crucial for tumor cell detachment and migration.

Addressing the question whether EMT is involved in tumor cell dissemination we studied the expression of EMT-related transcription factors. Although a tendency for a slight up-regulation of Snail was observed in CTC samples, we did not find convincing evidence that EMT is critically related to the process of tumor cell dissemination in the present mouse model. E-cadherin down-regulation in CTCs might be achieved through other mechanisms than EMT and since no up-regulation of mesenchymal genes was detected, the contribution of mesenchymal transition to tumor cell dissemination in this model remains uncertain. Interestingly, a tendency for the down-regulation of EGFR was found in CTCs of the mouse model. Abolished expression of EGFR was also seen in the DTCs from human patients. This is in contrast to a study by Wang et al. [183]. They report the mutual regulation of E-cadherin and EGFR. Loss of E-cadherin was found to up-regulate EGFR expression and enhanced the proliferation of squamous cell carcinoma of the head and neck in cell culture experiments. However, the influence of the microenvironment and the loss of close cell-cell contacts after tumor cell dissemination might additionally affect EGFR expression. Further studies are required to elucidate transcriptional regulation of EGFR during the dissemination process.

CTC survival and persistence in the circulation may require cellular strategies to develop resistance to apoptosis and anoikis. The up-regulation of the anti-apoptotic Bcl-2 protein in CTCs has been already reported in a previous study [181]. We were able to show a significant increase in the expression of Survivin, an inhibitor of apoptosis, in CTC specimens. Although p53 and β-catenin are linked to Survivin regulation, the expression of these genes does not seem to be differentially regulated in CTCs compared to solid tumor tissue. The involved regulatory pathways remain therefore elusive. Of the metastasis-associated genes included in the present study, we found CD44, a hyaluronic acid receptor, up-regulated in CTCs. Hyaluronic acid is the major component of the extracellular matrix and over-expression of CD44 in CTCs might be involved in homing of CTCs to distant organs [184]. However, increased expression was only seen for the standard isoform CD44s. The transcription level of the putatively

metastasis-associated extended isoform CD44v6 seemed not to be altered. Kuhn et al. [130] postulated that a complex of CD44v6, CO-029, CLDN7 and EpCAM contributes to metastasis formation. According to our results which do not indicate the differential expression of these genes in our experiments, a correlation could not be confirmed. This, however, might be attributed to the used HCT116-P cells, that generally do not seem to express the tetraspanin CO-029 which might be important for complex formation. Further studies with other cell lines might be useful to assess the role of the CD44v6/CO-029/CLDN7/EpCAM complex in tumor cell dissemination.

Conclusion and outlook

In conclusion, our approaches provide the opportunity to characterize CTCs and DTCs on the molecular level. We showed that the molecular profiles of CTCs and DTCs are not necessarily the same than in the primary tumor or the site of metastasis. Therefore, genomic analyses of cancer tissue biopsies for treatment decision might not be sufficient for systemic cancer therapy. Moreover, analyzed CTCs appeared to be very heterogeneous in general. Therefore, we believe that various subpopulations of CTCs exist and only a minority is able to induce metastatic growth. However, until today the metastasis-initiating capacity of human CTCs has never been proofed under experimental conditions. Neither the molecular traits nor the appropriate culturing conditions that could define the true metastatic precursors have been revealed so far. To obtain better insights into the biology of metastasis and the resistance to established therapies extended studies on CTCs/DTCs are required. Single cell research provides the means to gain comprehensive knowledge about individual cancer cell heterogeneity. Moreover, the direct characterization of single CTCs/DTCs may allow the identification of novel therapeutic targets to defeat systemic cancer disease and may elucidate the role of cancer stem cells in metastasis formation.

Material and methods

Methods

Blood samples

Peripheral blood samples were drawn from healthy donors from the *antecubital* vein and collected in EDTA tubes. To avoid epithelial cell contamination from skin puncture, the first 3 ml of blood were discarded. Patients' blood samples, from patients undergoing surgical therapy at the Department of surgery of the University Hospital Heidelberg, were obtained after induction of general anesthesia through a central venous catheter or from portal or liver vein. After collection, blood samples were immediately processed. The study was approved by the ethics committee of the University Heidelberg.

Bone marrow samples

Bone marrow samples were obtained after induction of general anesthesia by *iliac crest* biopsy.

Mononuclear cell collection

Blood samples were layered over 15 ml of LSM 1077 lymphocyte gradient. The samples were centrifuged at 4° C for 30 min at 300g without brake. The interphase containing the PBMC fraction was collected and spun down for 5 min at 300 g and washed once with PBS.

CTC/DTC enrichment and staining

PBMC were collected from about 27 ml blood and resuspended in 1.5 ml of beads buffer (PBS, 2 mM EDTA, 0.1% AB serum). After washing in beads buffer 450 µl magnetic beads were added to PBMCs and incubated in an overhead shaker for 20 min at 4° C. According to manufacturers instruction leukocyte depletion ought to be performed in a larger volume of buffer. Therefore, cell suspension was diluted with additional 8 ml of beads buffer before magnetic separation. Finally, the remaining supernatant was centrifuged 5 min 300g to obtain the CTC-enriched cell fraction. Pellet was resuspended in 150 µl of PBS buffer containing 10% FCS and 50 U/ml penicillin/ streptomycin. Per sample 5µl of Alexa-Fluor488-anti-EpCAM antibody was added.

After 20 min of incubation on ice the stained CTCs were identified under a fluorescence microscope. A similar protocol was applied for the bone marrow samples. In brief, bone marrow aspirates were collected according the mononuclear cell collection protocol. The further enrichment and staining of cell samples including potential DTCs was performed in the same way as described for the PBMC specimens. In case of CTC/DTC detection the identified cells were directly picked with the micromanipulator into RNA lysis buffer.

Cancer tissue specimens

Cancer tissue specimens were minced through a 45 µm cell strainer and collected in PBS. After centrifugation for 5 min and 300g the cell pellet was resuspended in 150 µl of PBS buffer containing 10% FCS and 50 U/ml penicillin/ streptomycin and stained for cancer cell identification with 5 µl anti-EpCAM Alexa488 antibody. Single cancer cells or small cell clusters were picked with the micromanipulator directly into RNA lysis buffer.

CellSearch^{TM} (Veridex)

7.5 ml of patients' blood samples, collected in CellSave^{TM} tubes, were used for CTC analysis according to the manufactures instructions.

Flow cytometry

Cell sorting was performed on FACS Aria II Flow Cytometer (Beckton Dickinson). Protein expression analysis by flow cytometry was carried out on FACS Calibur (Beckton Dickinson). Data was analyzed by FlowJo or FACS express software.

Immunohistochemistry

1-2 µm sections of formalin-fixed, paraffin-embedded tissue samples were mounted on object slides. After de-waxing in xylol and ethanol, antigen retrieval was achieved by incubating in target retrieval solution (Dako # S1699) for 20 min at 95° C. After washing in PBS, slides were incubated in methanol containing 0.3% H_2O_2 followed by washing in PBS. Non-specific binging sites were blocked with 20% normal goat serum for 1 hour. Slides were incubated with pre-labeled (Solulink All in one HRP conjugation kit #A-9002-001) Moc31-HRP antibody (IQ Products) over night. After

washing in PBS the immunoreaction was visualized by using ABC Pro Vecta Stain kit (Dako) according to the manual.

Cell culture

Cells were cultured at 37° C, 5% CO_2 and 95% humidity in RPMI supplemented with 10% FCS, 50 U/ml penicillin and 50 U/ml streptomycin in culture flasks. Cells were collected with trypsin EDTA and centrifugation for 5 min at 300 g.

RNA extraction

Total RNA was extracted by PicoPure RNA isolation kit (Arcturus) according to the manual.

RNA amplification

Amplification of total RNA was achieved by the use of WT Ovation RNA amplification kit (NuGen).

RT-PCR

Real time PCR using SYBR green was performed with LightCycler™ (Roche) according to the manual. An annealing temperature of 60° C and 55 cycles of amplification were used. Amplification plots were analyzed using the second derivative maximum method. All Primers were designed with Primer3Plus software and span large intronic regions to exclude amplification of genomic DNA.

gDNA amplification of single cells

Single cells were transferred into PCR tubes containing 9µl of PBS. 1µl of Proteinase K mix (1x OnePhorAll Buffer (OPA), 0.65% Tween, 0.65% Igepal, 1.3 mg/ml Proteinase K) was added. Cells were digested for 10 h at 42° C in a thermocycler followed by an inactivation step of 10 min, 80° C. For MseI digestion 0.4µl of MseI enzyme (high concentration 10.000 U/ml) was added. After an incubation time of 3 h at 37° C the enzyme was inactivated by heating the reaction mix to 65° C for 5 min. LIB1 (LIB1: AGTGGGATTCCTGCTGTCAGT) and ddMse11 (ddMSE11: TAACTGACAGCdd) primers were annealed in a reaction mix of 5µl 10x OPA , 5µl LIB1 100µM, 5µl ddMse11 100µM, 15µl H_2O in a water bath. For ligation 6µl of pre-annealed adapters combined with 2µl of 10mM ATP together with 2µl T4

DNA ligase were added to the digested cell. Ligation occurred over night at 15° C. Genomic DNA was amplified with the help of the Expand long template PCR system (Roche) in buffer 1. Therefore, 3µl Buffer 1, 2µl dNTPs 10µM, 25µl H_2O and 1µl DNA-polymerase mix were combined and added to the reaction mix. Amplification was performed in a thermocycler under the following conditions.

cycles	1x		14x		8x		22x		1x	
program	68°C	3min	94° C	40 s	64° C	40 s	94° C	40 s	68° C	220 s
			57° C	30 s	57° C	30 s + 1° C/cycle	65° C	30 s	4° C	∞
			68° C	90s + 1 s/Cycle	68° C	150 s + 1 s/Cycle	68° C	113 s + 1 s/Cycle		

For quality control of gDNA amplification a multiplex PCR was used. Therefore 0.3µl template was combined with 5µl Dream Taq Green PCR Mastermix (Fermentas), 1µl of 10 x primer mix (see table below) and 3.7 µl H_2O. PCR was run in thermocycler (pre-incubation 2 min 95° C; 33 x amplification 30 s 95° C / 40 s 60° C / 60 s 72° C; final elongation 5 min 72° C). Successful amplification was assessed on an agarose gel.

10x Primer mix:

Primer	final conc.	PCR Product
LAMC1 for	0.2 µM	111 bp
LAMC1 rev	0.2 µM	
GRIK5 for	0.1 µM	232 bp
GRIK5 rev	0.1 µM	
NRK9 for	0.1 µM	288 bp
NEK9 rev	0.1 µM	
CAPS for	0.08 µM	175 bp
CAPS rev	0.08 µM	
PICK1 for	0.08 µM	358 bp
PICK1 rev	0.08 µM	
DNAH9 for	0.08 µM	401 bp
DNAH9 rev	0.08 µM	

Gel electrophoresis of DNA using agarose gels

PCR products were verified for correct size by electrophoresis at RT using 1.5 % (w/v) agarose gels suspended in 1x TAE buffer. GelRed was used in the agarose gel to visualize the DNA under UV light.

Purification of PCR products

PCR clean up kit (Sigma) was used for purification of PCR products that were used for down stream applications such as sequencing, MSI analysis or CGH.

Sequencing

PCR products that were used for sequence analysis were generated by AmpliTaq Gold DNA Polymerase (Applied Biosystems). DNA was sequenced through sequencing laboratory (GATC / Konstanz).

Macrodissection and DNA extraction

For the macro-dissection procedure fresh frozen cancer tissue samples were used. Sections of 10µm were cut in a cryostat at -20° C and adhered to object slides. One section of each macrodissected sample was hematoxylin / eosin stained and examined by light microscopy to discriminate cancer tissue from stroma and healthy tissue. Between ten and 20 sections, depending on the amount of cancer tissue, were collected for gDNA extraction with the help of Qiagen DNeasy tissue kit according to manufacturers instructions. The estimated amount of cancer tissue within the dissected area was in every case at least 70 %.

Comparative genomic hybridization (CGH)

Array-CGH was performed externally by the group of Prof. Dr. N. Stöcklein at the University clinic Düsseldorf. For single cell analysis Agilent SurePrint G3 Human CGH 4x180K Oligo Microarray Kit (Cat. G4449A) was used. Amplified gDNA of single CTCs and matched leukocytes was hybridized against each other. Array-CGH analysis of macro-dissected snap frozen cancer tissue probed against matched healthy tissue was carried out with Agilent SurePrint G3 Human CGH 8x60K, Oligo Microarray Kit (Cat.: G4450A). Data processing was performed with the following software provided by Agilent Technologies: Scan Control V 7.0.3; Feature Extraction V 10.1.1.1 and Genomic Workbench V 5.0.14. The following parameters were used for data analysis:

Genome build: hg18;

Evaluation algorithm ADM-2, threshold 6;

Aberration filter: minimal number of probes in area 3, minimal mean log 2 ratio for area 0.25

MSI

Analyses of microsatellite instability were performed externally by the group of Dr. M. Kloor in the department of pathology at the University Hospital Heidelberg. Amplified gDNA of CTC samples as well as tumor samples were tested for MSI in NR21, NR24 and BAT 25 from the standard NCI/ICG-HNPCC marker panel [185] and analyzed as described previously [186].Genome amplification of CTC-derived gDNA led to disruption of the DNA while the afore mentioned markers remained unaffected and were selected for MSI analysis.

Mice

For experiments BalbC nu/nu and NOD/SCID gamma mice were used. Authorization number 35-9185.81/G-7/10 was obtained from the national authorities for research experiments on animals. Mice were kept under specific pathogen free conditions in IBF animal facility Heidelberg.

Orthotopic injection of tumor cells

Mice were anesthetized by 3-5% of isofluran. A small abdominal incision was made and the cecum was exteriorized. 10^5 tumor cells in a total volume of 20µl of Matigel (10mg/ml) were injected orthotopically into the coecal wall by microinjection. Incisions were closed with 6.0 resorbable sutures. For postoperative pain relief 4mg/kg Rimadyl was given subcutaneously.

Statistical evaluation

To identify independent factors associated with detection of CTCs in the CVBC and MVBC the following statistical tests were applied. Categorical data were presented as absolute and relative frequencies and compared using the χ^2-test. Continuous data were presented as median and range and compared using the Wilcoxon test. Furthermore, the mean number of CTCs was reported. Analysis of covariance (ANOVA) was applied to compare continuous data of more than two groups.

The significance of mRNA expression data differences was evaluated with Student's T-test. P – values below 0.05 were considered as statistical significant.

Material

Equipment

CellSearch (Veridex)	Ortho Clinical Diagnostics
Centrifuges	Heraeus Multifuge 3; Eppendorf Centrifuge 5810R
Electrophoresis power supply	Consort E835
Electrophoresis units	Neolab
FACS Aria II	BD
FACS Calibur	BD
Hood	Heraeus / Herasafe
Incubator	Binder
Light cycler 480	Roche
Microinjection system	World Precision Instruments
Microscopes	Leica
Nanodrop	Peqlab
PCR Thermocycler System	Eppendorf
pipettes	Gilson
Shakers	Assitent RM5
Thermomixer	Eppendorf
Transferman Micromanipulator	Eppendorf
UV- transilluminator	BioDoc-It System
Vortex	Scientific industries Vortex Genie 2

Chemicals and reagents

Agarose	Invitrogen
Antigen retrieval solution	Dako
ATP	NEB
Deoxyribonucleotide triphosphates (dNTPs)	NEB
Dimethylsulfoxide (DMSO)	Sigma
DMEM	PAA
DNA ladder	Fermentas
DNA loading dye	Fermentas
DynaBeads	Dynal / Invitrogen
EDTA	Roth
Ethanol	Roth
Fetal Calf Serum (FCS)	PAA
Gel red DNA stain	Biotium
H_2O_2 (30%)	Roth
Igepal-CA630	Sigma
Isofluran	CP Pharma
LSM 1077 lymphocyte	PAA
Matrigel	BD
Normal goat serum	Vector
Paraformaldehyde	Roth
PBS	PAA

Penicillin/ Streptomycin	PAA
RPMI	PAA
Trypsin EDTA	PAA
Tween	Sigma
Xylol	Roth

Kits

ABC Vecta Stain kit	Dako
Expand long template PCR system	Roche
LC 480 RT-PCR kit	Roche
PCR clean-up kit	Sigma
Pico Pure RNA extraction	Applied bioscience
Solu Link All in one HRP conjugation kit	Solu link
WT-Ovation RNA amplification kit	Nugen

Enzymes

AmpliTaq Gold	Applied biosystems
DreamTag Green	Fermentas
MseI	Roche
Proteinase K	Roche
T4 DNA ligase	Roche

Antibodies

Name of antibody	Target	company
EpCAM-AlexaFluor 488	EpCAM	Biolegend
CD45-PE	CD45	BD
EpCAM-K/S1-4	EpCAM	Santa Cruz
CD66	CEA	BD
Moc-31	EpCAM	IQ-Products
Ber-Ep4 EpCAM-FITC	EpCAM	Dako
Mouse IgG1k Con-FITC	Isotype control	BD
EpCAM-VU-1D9	EpCAM	ABR
mouse IgG1-PE isotype control	Isotype control	BD
HEA125	EpCAM	Gift of Mollberg

Oligonucleotides

All used oligonucleotides were obtained from Invitrogen.

Target	Name	sequence	Purpose
Ascl2	ASCL2_1220F	GAGGGGAGAGGATTTTCTAAGG	RT-PCR
Ascl2	ASCL2_1220R	TTATTACGCCCCAGGTCAAG	RT-PCR
Bax	Bax-Fwd	TCTGACGGCAACTTCAACTG	RT-PCR
Bax	Bax-Rev	GGAGGAAGTCCAATGTCCAG	RT-PCR
B-Catenin	Bcat_2344F	TTCCGAATGTCTGAGGACAAG	RT-PCR
B-Catenin	Bcat_2344R	TGGGCACCAATATCAAGTCC	RT-PCR

Bmi1	Bmi1_1209Fwd	GATACTTACGATGCCCAGCAG	RT-PCR
Bmi1	Bmi-1_1209Rev	GAAGTGGACCATTCCTTCTCC	RT-PCR
Calreticulin	Calr_1040F	CCACCCAGAAATTGACAACC	RT-PCR
Calreticulin	Calr_1040R	TGTCAAAGATGGTGCCAGAC	RT-PCR
CD133	CD133_2410Fa	AAAGTGGCATCGTGCAAAC	RT-PCR
CD133	CD133_2410Ra	CCGAATCCATTCGACGATAG	RT-PCR
CD151	CD151_759Fwd	AGCAACAACTCACAGGACTGG	RT-PCR
CD151	CD151_759Rev	TGCTCCTGGATGAAGGTCTC	RT-PCR
CD166	ALCAM3650Fwd	GAACACTGCACAGCGATTTC	RT-PCR
CD166	ALCAM3650Rev	CAAACACCAGTTTTCCTTTTCC	RT-PCR
CD26	CD26Fwd_2670	AGTCAGCTCAGATCTCCAAAGC	RT-PCR
CD26	CD26Rev_2670	TGTGCTGTGCTGCTAGCTATTC	RT-PCR
CD44s	CD44s_Fa	AAAGGAGCAGCACTTCAGGA	RT-PCR
CD44s	CD44s Ra	TGTGTCTTGGTCTCTGGTAGC	RT-PCR
CD44v6	CD44v6_Fb2	GTACAACGGAAGAAACAGCTACC	RT-PCR
CD44v6	CD44v6_Rb2	TGTTGTCGAATGGGAGTCTTC	RT-PCR
CD45	CD45_3819Fb	TCTCTTAGAAAGTGCGGAAACAG	RT-PCR
CD45	CD45_3819Rb	TCCATTCTGAGCAGGGTAGG	RT-PCR
CD47	CD47_1057F	ATAGCCTATATCCTCGCTGTGG	RT-PCR
CD47	CD47-1057R	CGGAGTCCATCACTTCACTTC	RT-PCR
CEA	CEA5 Fwd_2267	CATGATTGGAGTGCTGGTTG	RT-PCR
CEA	CEA5 Rev_2267	TAGGATGGTCTCGATCTCTGG	RT-PCR
CK18	CK18_1288Fwd	CCCTGCTGAACATCAAGGTC	RT-PCR
CK18	CK18_1288Rev	TCAGACACCACTTTGCCATC	RT-PCR
CK19	CK19TN_Fwd563	GCGAGCTAGAGGTGAAGATCC	RT-PCR
CK19	CK19TN_Rev563	TGTCGATCTGCAGGACAATC	RT-PCR
CK20	CK20 1219 Fa	ACGCCAGAACAACGAATACC	RT-PCR
CK20	CK20_1219 Ra	GCCATCCACTACTTCTTGCAC	RT-PCR
CLDN7	CLDN7_656Fa	TGAGCTGCAAAATGTACGACTC	RT-PCR
CLDN7	CLDN7_656 Ra	CACAAACATGGCCAGGAAG	RT-PCR
c-Myc	C-myc_1327F	TGCTCCATGAGGAGACACC	RT-PCR
c-Myc	C-myc_1327R	GATCCAGACTCTGACCTTTTGC	RT-PCR
CO029	CO-029_444Ra	ACAGCTCCTAGGATACCTGTCG	RT-PCR
CO-029	CO-029_444Fa	TGATTGCTGTAGGTGCCATC	RT-PCR
CXCR4	CXCR4_1000F	CATCATGGTTGGCCTTATCC	RT-PCR
CXCR4	CXCR4_1000R	CGATGCTGATCCCAATGTAG	RT-PCR
Cyclin D1	CCND1Fwd_933	TCCTCTCCAGAGTGATCAAGTG	RT-PCR
Cyclin D1	CCND1Rev_933	TTGGGGTCCATGTTCTGC	RT-PCR
DLG7	DLGPA5fwd_2338	TGAAAGCAGGAGCAGCATAG	RT-PCR
DLG7	DLGPA5rev_2338	ATCTGCTACTCCACCAGCAAG	RT-PCR
E-cadherin	E-Cad2563F	AGAGGACCAGGACTTTGACTTG	RT-PCR
E-cadherin	E-Cad2563R	TCAGTATCAGCCGCTTTCAG	RT-PCR
EGFR	EGFR_3517Fwd	TTCTTGCAGCGATACAGCTC	RT-PCR
EGFR	EGFR_3517Rev	TGGGAACGGACTGGTTTATG	RT-PCR
EGR1	EGR1_577F	CAGCACCTTCAACCCTCAG	RT-PCR
EGR1	EGR1_577R	AGCGGCCAGTATAGGTGATG	RT-PCR
EpCAM	EpCAM_733Fa	CTGGATCCAAAATTTATCACGAG	RT-PCR
EpCAM	EpCAM_733Ra	GTTCCCCATTTACTGTCAGGTC	RT-PCR
HDAC1	HDAC1_1042Fwd	TTAACCTGCCTATGCTGATGC	RT-PCR

HDAC1	HDAC1-1042Rev	GAAGGACTGATGTGGAGCTTG	RT-PCR
HDAC3	HDAC3_1283Fwd	ACAGGACTGATGAGGCTGATG	RT-PCR
HDAC3	HDAC3_1283Rev	CCCAACCAAGAGGTGAAAAG	RT-PCR
hFN	hFN1	CAGTGGGAGACCTCGAGAAG	RT-PCR
hFN	hFN1	TCCCTCGGAACATCAGAAAC	RT-PCR
Ki67	Ki67_8625Fwd	ACAAAAGGTGCTTGAGGTCTG	RT-PCR
Ki67	Ki67_8625Rev	CCTTTCCCTTTCTGATTCTGC	RT-PCR
Lgr5	LGR5_1685F	TCATTCAGTGCAGTGTTCACC	RT-PCR
Lgr5	LGR5_1685R	CTGCGATGACCCCAATTAAC	RT-PCR
Maspin	SerpinB5Fwd_877	ACACCAAACCAGTGCAGATG	RT-PCR
Maspin	SerpinB5Rev_877	CTGTGACAGTGACTCTGAGTTGAG	RT-PCR
MMP7	MMP7Fwd_822	TTGGGTATGGGACATTCCTC	RT-PCR
MMP7	MMP7Rev_822	GAATGGATGTTCTGCCTGAAG	RT-PCR
MRP2	MRP2_4453Fwd	AGCCTGCAACTTGGGTTATC	RT-PCR
MRP2	MRP2_4453Rev	TGGTCGTCTGAATGAGGTTG	RT-PCR
MRP4	MRP2_2655Fwd	CTGCCGCTGACGTTTTTAG	RT-PCR
MRP4	MRP2_2655Fwd	AATCCCAGCACTGAGGTTTG	RT-PCR
MRP5	MRP2_453Fwd	AGATGCCTTGGAAACAGCAG	RT-PCR
MRP5	MRP2_453Fwd	AAAAAGCCCAGCATTGTCC	RT-PCR
Musashi	Msi1Fwd_1174	CTTTGATTGCCACAGCCTTC	RT-PCR
Musashi	MsiRev_1174	GCTGGCTCACTCGTGGTC	RT-PCR
N-cadherin	N-Cad2973Fneu	GAATGGATGAAAGACCCATCC	RT-PCR
N-cadherin	N-Cad2973R	AGTCATATGGTGGAGCTGTGG	RT-PCR
Olfactomedin 4	Olfm4_730Fa	AGATCAAAACACCCCTGTCG	RT-PCR
Olfactomedin 4	Olfm4_730Ra	GAAAACCCTCTCCAGTTGAGC	RT-PCR
P21	p21Fwd_571	ATGTGGACCTGTCACTGTCTTG	RT-PCR
P21	p21Rev_571	GGATTAGGGCTTCCTCTTGG	RT-PCR
Snail1	SnailR_680	TCTTGACATCTGAGTGGGTCTG	RT-PCR
Snail1	SnailFb_680	TCTAGGCCCTGGCTGCTAC	RT-PCR
Survivin	Birc5Fwd_232	TTTTCATCGTCGTCCCTAGC	RT-PCR
Survivin	Birc5Rev_232	AGCCCGGATGATACAAACAG	RT-PCR
TP53	p53Fwd_1297	TGAATGAGGCCTTGGAACTC	RT-PCR
TP53	P53Rev_1297	TTTTATGGCGGGAGGTAGAC	RT-PCR
Twist1	Twi1100Fb	GGGCCGGAGACCTAGATG	RT-PCR
Twist1	Twi1100R	CCACGCCCTGTTTCTTTG	RT-PCR
Vimentin	Vim1483F	TTTTCCTCCCTGAACCTGAG	RT-PCR
Vimentin	Vim1483R	CGTGATGCTGAGAAGTTTCG	RT-PCR
ZEB1	ZEB1_2664Fwd	CTAGCTGCCAATAAGCAAACG	RT-PCR
ZEB1	ZEB1_2664Rev	TTGGGCGGTGTAGAATCAG	RT-PCR
B-Actin	b-Act_1404Fa	ATGTGGCCGAGGACTTTGATT	RT-PCR
B-Actin	b-Act_1510Ra	AGTGGGGTGGCTTTTAGGATG	RT-PCR

Preamplification

P53E5/6	E5/6Fwdx	TTCACTTGTGCCCTGACTTTCAAC	Tissue
P53E5/6	E5/6Revx	GCCACTGACAACCACCCTTAAC	Tissue
P53E7	E789seqP53Fwd	CCTCATCTTGGGCCTGTGTTATC	Tissue
P53E7	E7 Revx	TGGAAGAAATCGGTAAGAGGTGG	Tissue
P53E8	E8Fwdx	GGACCTGATTTCCTTACTGCCTC	Tissue
P53E8	E8Revx	CTGAGGCATAACTGCACCCTTG	Tissue
BRAF	BRAF_Fwd	GTGGGATTCCTGCTGTCAGTTAAAG	CTC
BRAF	BRAF-Rev	AGCCTCAATTCTTACCATCCAC	CTC
Kras	KrasFwd	TCCTGCTGTCAGTTAACCTTATG	CTC
Kras	KrasRev	CTGTCAGTTAAAACAAGATTTACC	CTC
P53E5/6	P53Exon5/6Fwd	GCTGTCAGTTAATGTGTGATCTC	CTC
P53E5/6	P53 Exon5/6Rev	GTCAGTTAACCCCTCCTCCCAG	CTC
P53E7	P53E7Fwd	GCAGTGGCTCATGCCTGTAATC	CTC
P53E7	P53E7Rev	GAGTGGGAGCAGTAAGGAGA	CTC
P53E8	P53E8_Fwd	GGGATTCCTGCTGTCAGTTAAATGG	CTC
P53E8	P53E8 _Rev	CATTGTCTTTGAGGCATC	CTC

Sequencing

Target	Name	sequence	Purpose
BRAF	BRAF fwd	CTCTTCATAATGCTTGCTC	CTC + Tissue
BRAF	BRAF rev	CCATCCACAAAATGGATCC	CTC + Tissue
Kras	Kras Fwd	TCCTGCTGTCAGTTAACC	CTC
Kras	Kras Rev1	ACCTCTATTGTTGGATC	CTC
P53E5/6	p53E5/E6Fwd	GTGATCTCTGACTCCTGTC	CTC
P53E5/6	p53 rev1	AGAGACCCCAGTTGCAAAC	CTC
P53 E7	p53E7rev	GGGAGCAGTAAGGAGATTC	CTC
P53 E8	p53E8-Fwd	AAATGGGACAGGTAGGAC	CTC
P53 E8	p53E8-Rev	CATTGTCTTTGAGGCATC	CTC
P53E5	E5p53gDNAFwd	TTCACTTGTGCCCTGAC	Tissue
P53E5	E5p53gDNARev	ACTGACAACCACCCTTAAC	Tissue
P53E7	E7p53gDNAFwd	ATCTTGGGCCTGTGTTATC	Tissue
P53E7	E7p53gDNARev	TGGAAGAAATCGGTAAGAG	Tissue
P53E8	E8p53gDNaFwd	GGACCTGATTTCCTTACTG	Tissue
P53E8	E8p53gDNARev	CTGAGGCATAACTGCAC	Tissue
Kras	KrasFwd	AAGGCCTGCTGAAAATGACTG	Tissue
Kras	KrasRev	AGAATGGTCCTGCACCAGTAA	Tissue

MSI Analysis

Target	Name	sequence	Purpose
NR21	NR-21_fwd	TAAATGTATGTCTCCCCTGG	MSI analysis
NR21	NR-21_rev	ATTCCTACTCCGCATTCACAA	MSI analysis
NR24	NR-24_fwd	CCATTGCTGAATTTTACCTC	MSI analysis
NR24	NR-24_rev	ATTGTGCCATTGCATTCCAAA	MSI analysis
BAT25	Bat25F	TCGCCTCCAAGAATGTAAGT	MSI analysis
BAT25	Bat25R	TATGGCTCTAAAATGCTCTGTTC	MSI analysis

Multiplex Primer quality control CTC amplification

LAMC1For	TCTGCTTTGGGCATTCTTCT
LAMC1Rev	TTCTAACAGGTTGGGGGATG
CADPSFor	CCCCACCCTTCTTCACTACA
CADPSRev	GTGTGCACATACCACCGAAG
GRIK5For	CTAGCTCCCACCAACCTCAG
GRIKRev	CTCGATGATCCCGTTGATCT
NEK9For	GCAGGAGGGAACCTGTATGA
NEK9Rev	CAGGAAAGAAAGCCCACAGA
PICK1For	TCGTATGCTGGAGTCCTGTG
PICK1Rev	GGGATGGCTTTGTTGAGGTA
DNAH9For	GGGTCTCATCACCAGCATTT
DNAH9Rev	GCCATCTTCCACATGGTCTT

Cell lines

Cell lines were obtained from DSMZ and their identity is regularly confirmed.

- HCT116
- HT29
- SW480
- DLD1
- Colo205

Appendix

List of Figures

List of tables

List of abbreviations

AS-CTC	Ascites-asscociated CTC
ATP	Adenosine triphosphate
BMDC	Bone marrow derived cells
CGH	Comparative genomic hybridization
CRC	Colorectal cancer
CTC	Circulating tumor cell
CVBC	Central venous blood compartment
DNA	Desoxyribonucleic acid
DTC	Disseminated tumor cell
gDNA	genomic DNA
FACS	Fluorescence activated cell sorting
mRNA	Messenger RNA
MS	Microsatellite
MSI	Microsatellite instability
MSS	Microsatellite stable
MVBC	Mesenterial venous blood compartment
NCI	National cancer institute
NSCLC	Non small cell lung cancer
PBMC	Peripheral blood mononucleated cells
PCR	Polymerase chain reaction
ptci	post tumor cell injection
RNA	Ribonucleic acid
Wt	Wild type

Supplementary figure 1: Genomic analyses of CTCs. MSS= microsatellite stable; MSI=microsatellite instablity; wt =wild type; mutations are indicated in red; number in brackets indicate the number of analyzed cells

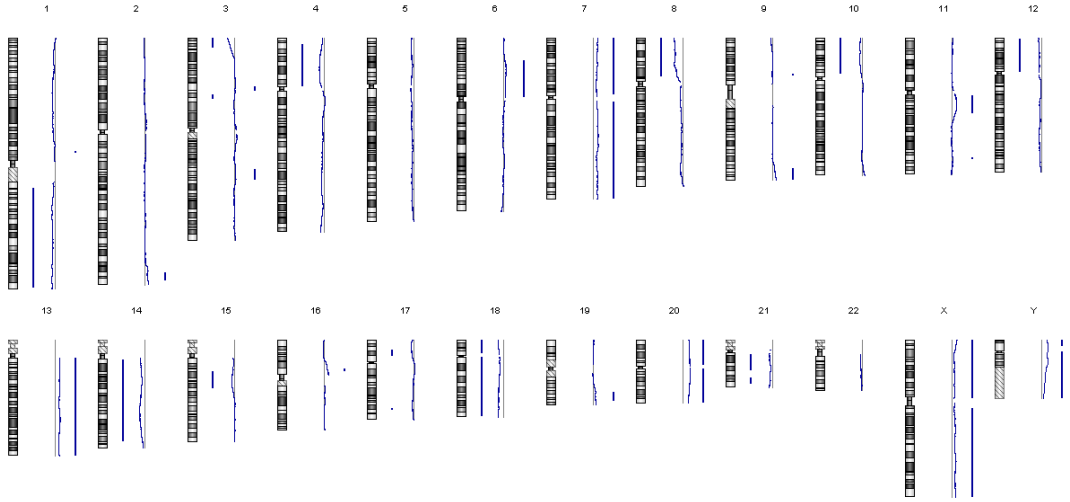
	MSS / MSI		Kras		Braf		TP53 Exon 5/6		TP53 Exon 7		TP53 Exon8	
Patient ID	**CTC**	**tissue**	**CTC**	**tissue**	**CTC**	**tissue**	**CTC**	**tissue**	**CTC**	**tissue**	**CTC**	**tissue**
HD1960	MSS (1)	MSS	X	wt	wt (1)	wt	X	wt	X	wt	X	wt
HD2026	MSS (1)	MSS	X	wt	wt (1)	wt	wt (1)	wt	wt (1)	wt	X	FS_C184
HD2030	MSS (5)	X	wt (8)	wt	wt (4)	wt	wt (5)	wt	wt (7)	wt	wt (6)	wt
HD2050	MSS (1)	X	wt (1)	X	wt (1)	X	wt (1)	X	wt (1)	X	wt (1)	X
HD2054	MSS (3)	MSS	wt (2)	wt	V600E(1)/ wt(1)	V600E	wt (2)	wt	wt (2)	wt	wt (2)	wt
HD2068	MSS (2)	MSS	wt (1)	wt	X	wt	R175H (2)	R175H	wt (1)	wt	wt (1)	wt
HD2077	MSS (1)	MSS	wt (1)	wt	X	wt	wt (1)	wt	wt (1)	wt	wt (1)	wt
HD2078	MSS (1)	MSS	wt (1)	wt	wt (1)	wt	wt (1)	wt	X	X	wt (1)	X
HD2091	MSS (1)	MSS	wt (1)	wt	X	wt	wt (1)	wt	wt (1)	wt	R273C (1)	R273C
HD2094	MSS (1)	MSS	X	wt	Wt	wt	wt (1)	X	wt (1)	X	wt (1)	X
HD2095	MSS (13)	X	wt (8)	G12D	wt (1)	wt	wt (9)	wt	wt (9)	wt	wt (7)	wt
HD2101	MSS (1)	MSS	wt (1)	G12D	X	wt	wt (1)	wt	wt (1)	wt	wt (1)	wt
HD2113	MSS (4)	MSS	wt (1)	wt	wt (2)	wt	wt (2)	wt	wt (4)	wt	R282W (4)	R282W
HD2126	MSS (1)	MSS	X	G12D	wt (1)	wt	A159V (1)	A159V	wt (1)	wt	wt (1)	wt
HD2131	MSS (1)	MSS	wt (1)	G12V	wt (1)	wt	X	wt	wt (1)	wt	wt (1)	wt
HD2165	MSS (14)	MSS	wt (8L; 5P)	wt	wt (8)	wt	wt (13)	wt	wt (14)	wt	wt (14)	wt
HD2179	MSS (1)	MSS	wt (1)	wt	wt (2)	V600E	R175H (1)	R175H	X	wt	wt (1)	wt
HD2201	MSS (1)	MSS	G12C (1)	G12C	X	wt	R175H (1)	R175H	X	wt	wt (2)	wt
HD2202	MSS (1)	MSS	G13D (1)	wt	wt (1)	wt	wt (1)	wt	wt (1)	wt	wt (1)	wt
HD2203	MSS (3)	X	wt (1)	X	wt (3)	wt	wt (2)	wt	wt (2)	wt	EE285, 286DA(3)	wt

HD2215	MSS(9)/ MSI(2)	MSS	G12S(4)/ wt (2)	G12S	wt (6)	wt	wt (2)	wt	wt (4)	wt	wt (5)	wt
HD2224	MSS (7)	MSS	wt (7)	G12R	wt (6)	wt	wt (6)	wt	wt (5)	wt	wt (7)	wt
HD2231	MSS (4)	MSS	wt (4)	wt	wt (4)	X	wt (4)	wt	G245S (3)	G245S	wt (4)	wt
HD2246	MSS (3)	X	wt (3)	xx	wt (3)	X	wt (3)	X	wt (3)	X	wt (2)	X
HD2257	X	MSS	X	wt	wt (1)	X	X	wt	wt (1)	R248Q	wt (1)	wt
HD2288	MSS (7)	MSS	G12A(4)/ wt(2)	G12A	wt (5)	wt	wt (5)	wt	wt (6)	wt	R273C(4)/ wt(1)	R273C
HD2295	MSS (7)	MSS	wt (7)	wt	wt (6)	V600E	wt (7)	wt	wt (7)	wt	wt (8)	wt
HD2328	MSS (14)	MSS	wt (12)	wt	wt (9)	wt	F134C(3L)+ wt(3LV,4M es)	F134C_LM; wt_Tu?	wt (10)	wt	wt (11)	wt
HD2334	MSS (8)	MSS	G12S(7)/ wt(1)	G12S	wt (8)	X	wt (7)	X	wt (7)	wt	wt (8)	wt
HD2341	MSI(2) /MSS(1)	X	G13D(2)/ wt(1)	X	wt (2)	X	wt (3)	X	wt (3)	X	wt (3)	X
HD2351	MSS (1)	MSS	wt(1)	X	wt (1)	X	wt (1)	X	wt (1)	wt	wt (1)	

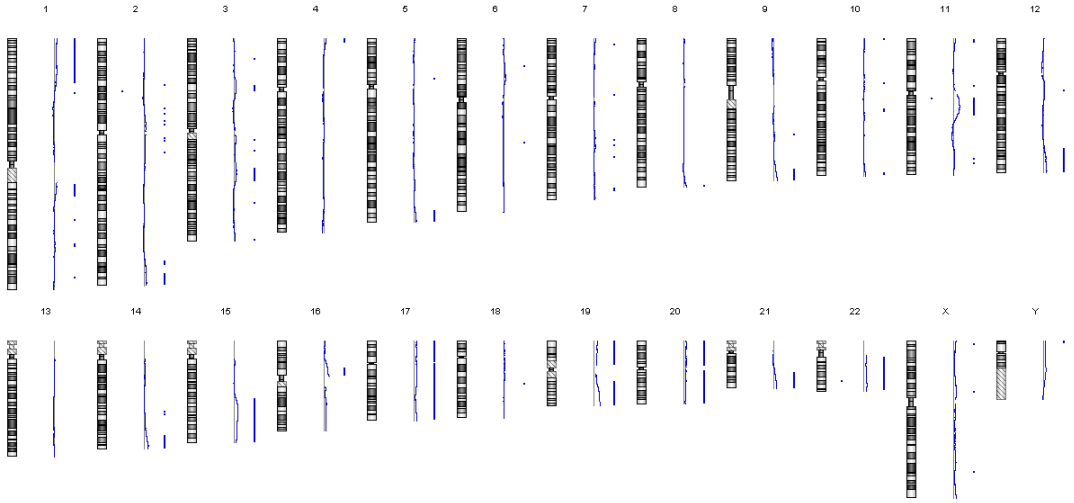
Supplementary figure 2: Patients characteristics. Patients from whom CTCs and cancer tissue specimens were obtained for genomic analyses are listed. f=female, m=male, Tu=tumor, LM=liver metastasis, LuMet=lung metastasis, LN=lymph node metastasis, PB peripheral blood, LV= liver vein, PV=portal vein

Patient ID	Sex	Age	UICC Stage	Analyzed tissue	CTC obtained from	Detected genomic disparity between CTCs and cancer tissue
HD1960	f	81	IV	Tu	PB	
HD2026	M	66	IV	LM	PB	
HD2030	m	54	IV	Tu	LV	
HD2050	M	53	I	-	PV	
HD2054	m	53	III	Tu	PB	
HD2068	m	36	IV	LM+LuMet	LV	
HD2077	f	72	I	Tu1+Tu2	LV	
HD2078	m	68	IV	Tu	LV	
HD2091	M	53	IV	Tu	LV	
HD2094	f	74	IV	LM	LV	
HD2095	f	69	IV	Tu	LV	x
HD2101	f	85	III	Tu	LV	x
HD2113	f	48	IV	Tu	LV	
HD2126	f	64	IV	Tu	LV	
HD2131	f	61	IV	Tu	PB	x
HD2165	m	50	III	LM	LV+PV	
HD2179	f	60	IV	LM	PB	x
HD2201	m	65	IV	LM	LV+PV	
HD2202	m	56	III	LN-Met	PV	x
HD2203	f	38	IV	-	PB	
HD2215	f	85	IV	LM	LV	
HD2224	m	62	III	LN-Met	LV	x
HD2231	m	63	IV	LM	LV	
HD2246	m	64	I	-	LV	
HD2257	f	55	IV	LM	PB	x
HD2288	F	59	IV	LM	LV	
HD2295	f	80	III	Tu	LV	x
HD2328	f	64	IV	Tu	LV+PV	
HD2334	m	74	III	LN-Met	PV	
HD2341	m	51	IV	-	PV	
HD2351	w	83	II	Tu	PV	

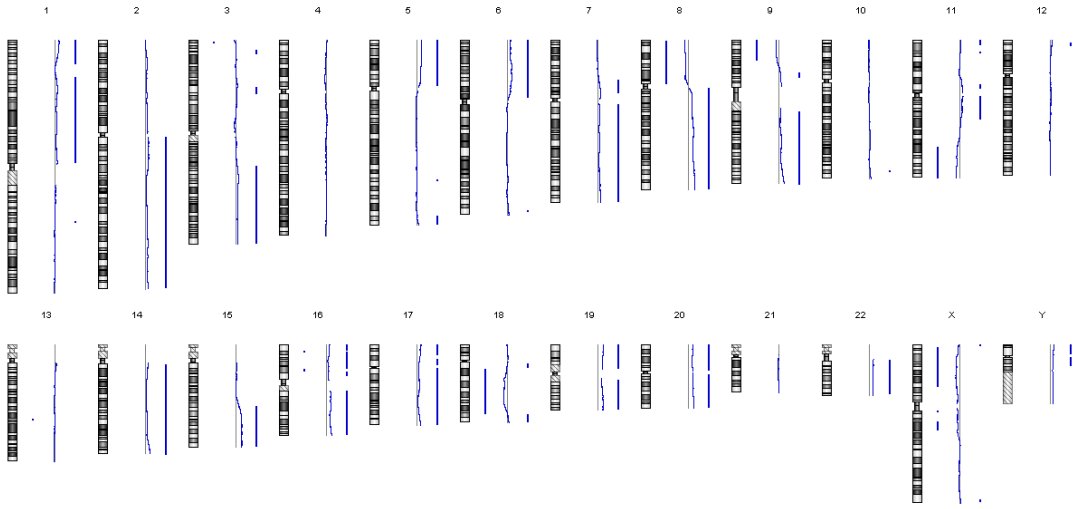
Supplementary figure 3: Array-CGH profiles of CTCs and cancer tissue samples. The human set of chromosomes is shown (1-22, X, Y). Beside each chromosome the detected signal intensity of the array-CGH analysis is depicted. Amplitudes to the left reflect chromosomal losses. Amplitudes to the right reflect chromosomal gains. CTCs were obtained from blood samples as indicated: PB=peripheral or central blood, LV=liver vein, PV=portal vein.



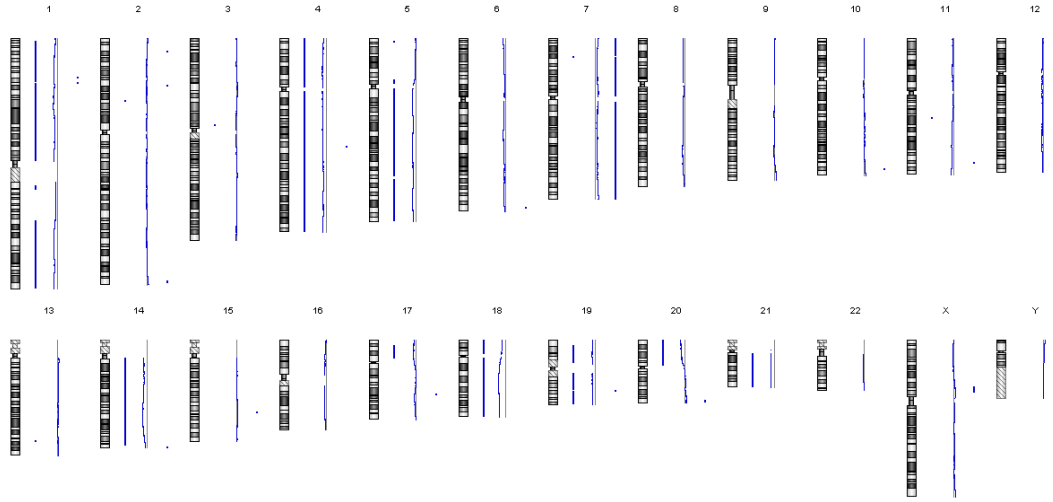
Matrix-CGH profile HD2026 CTC (PB)



Matrix-CGH profile HD2030 CTC (LV)

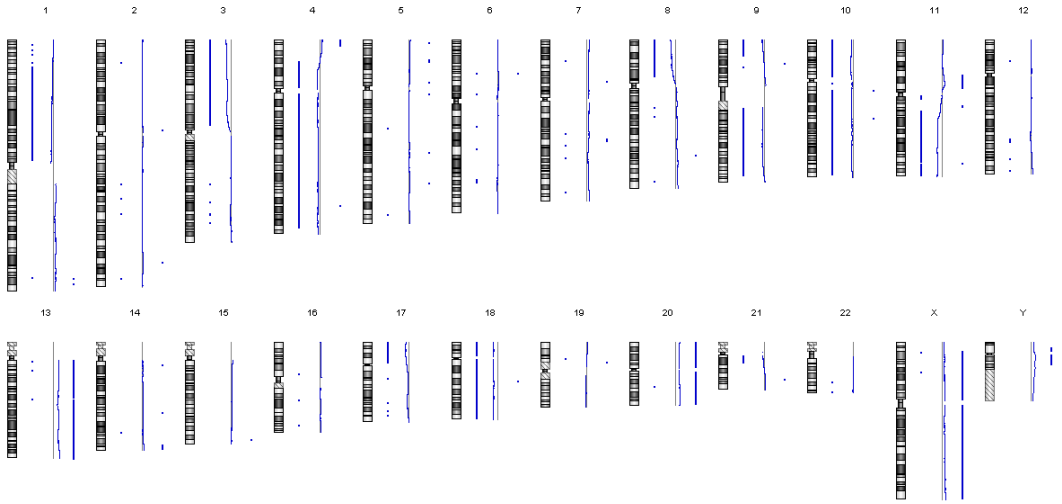


Matrix-CGH profile HD2095 CTC (LV)

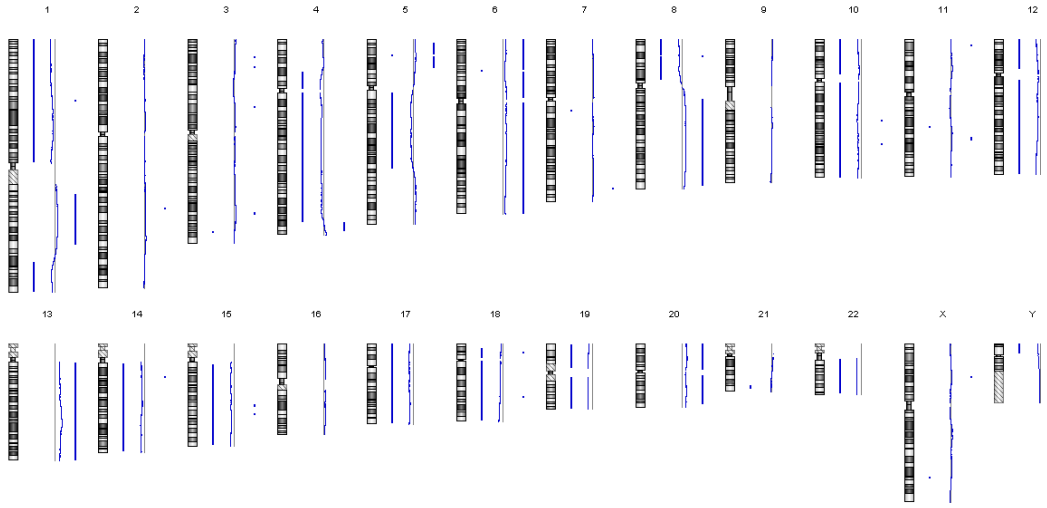


Matrix-CGH profile HD 2126 CTC (LV)

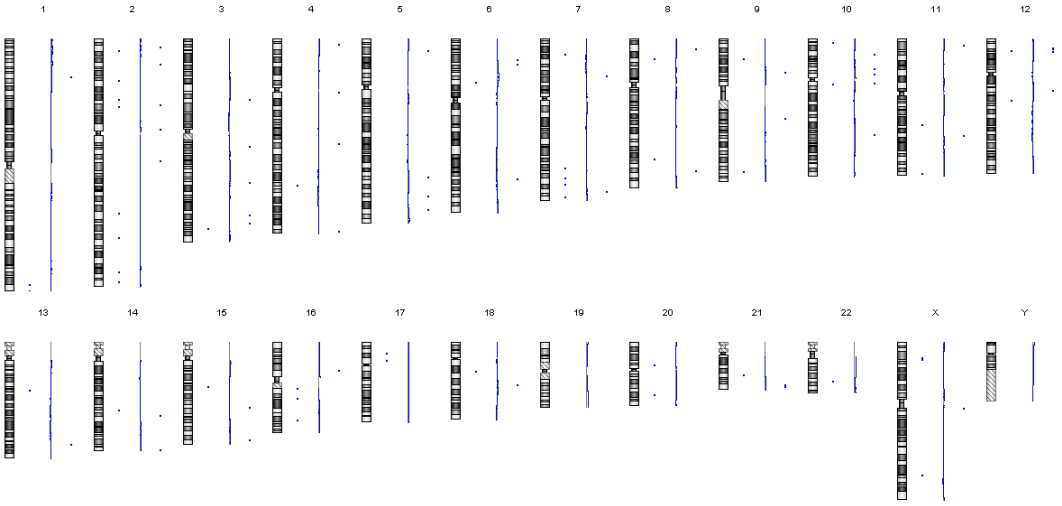
Matrix-CGH profile HD 2224 CTC (LV)



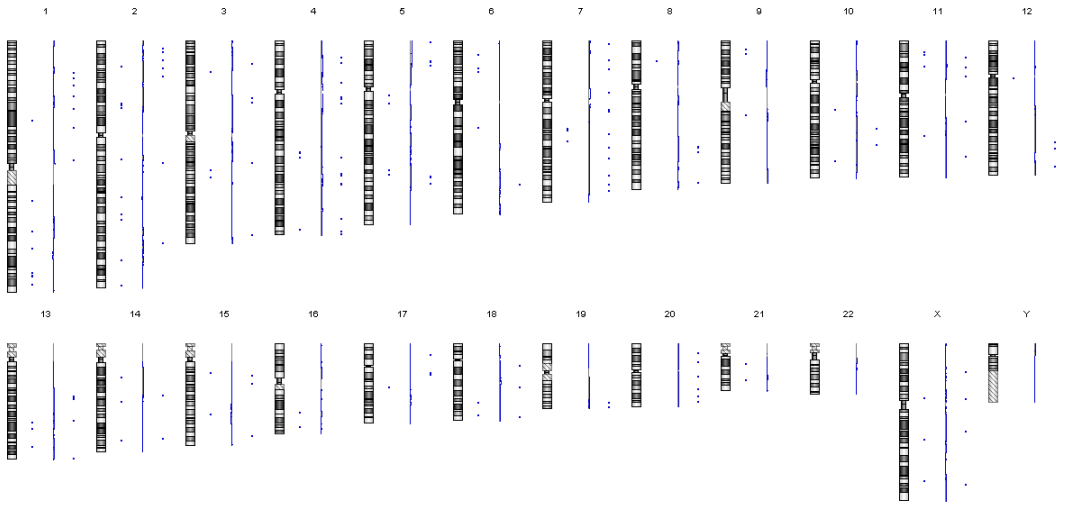
Matrix-CGH profile HD2231 CTC (LV)



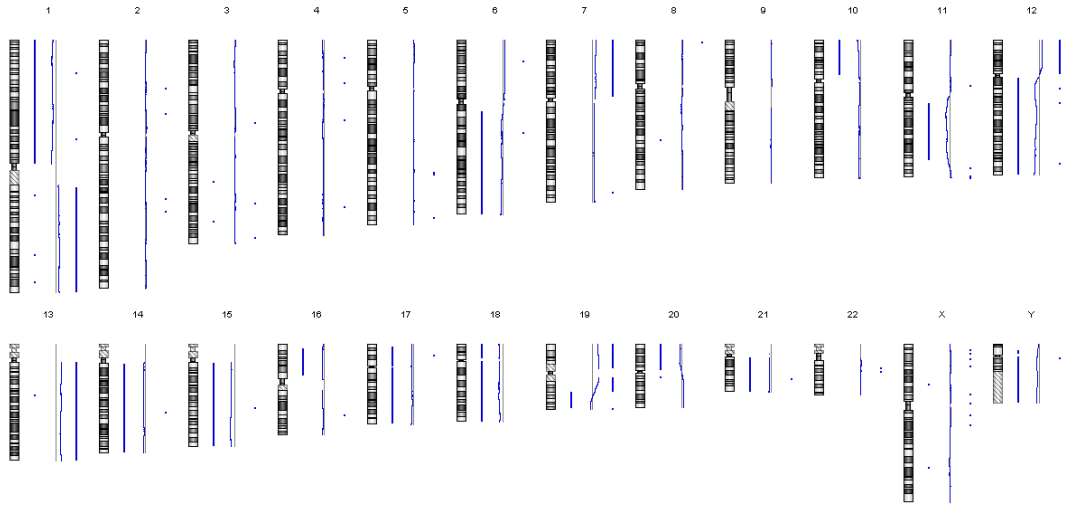
Matrix-CGH profile HD2288 CTC (LV)



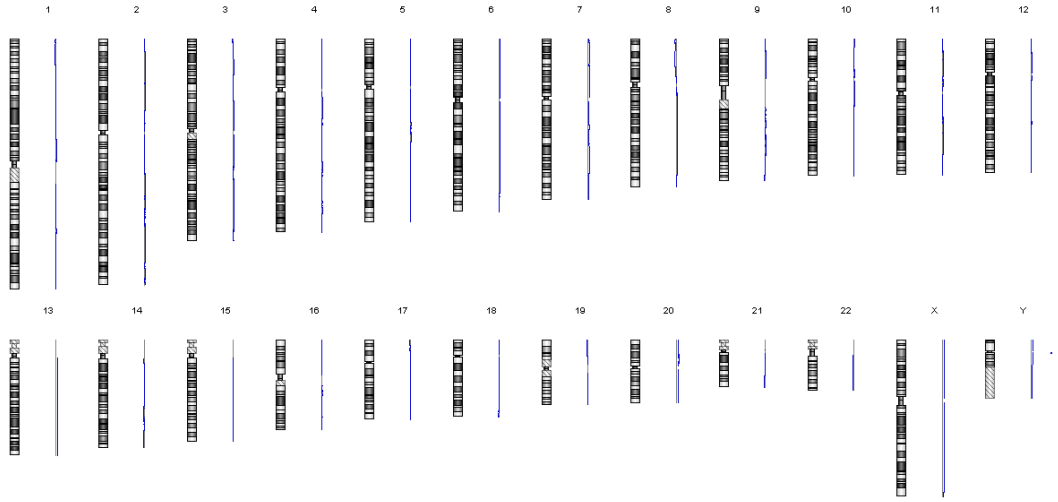
Matrix-CGH profile HD2295 CTC (LV)



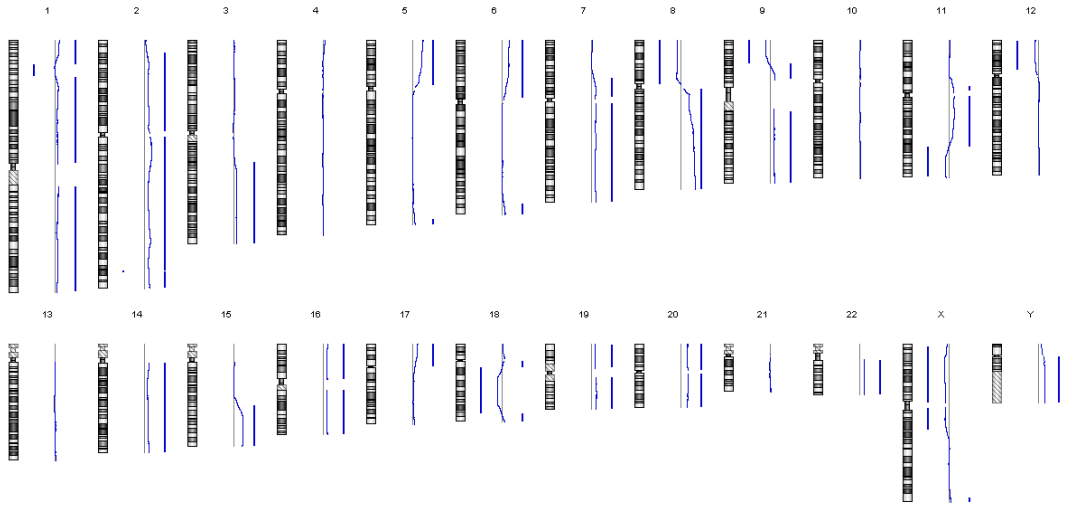
Matrix-CGH profile HD2328 CTC (LV)



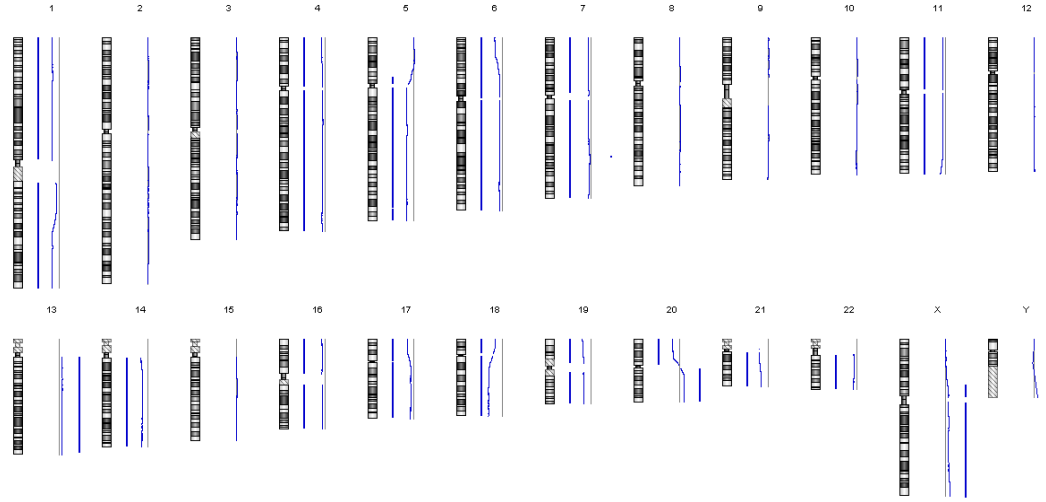
Matrix-CGH profile HD2334 CTC (PV)



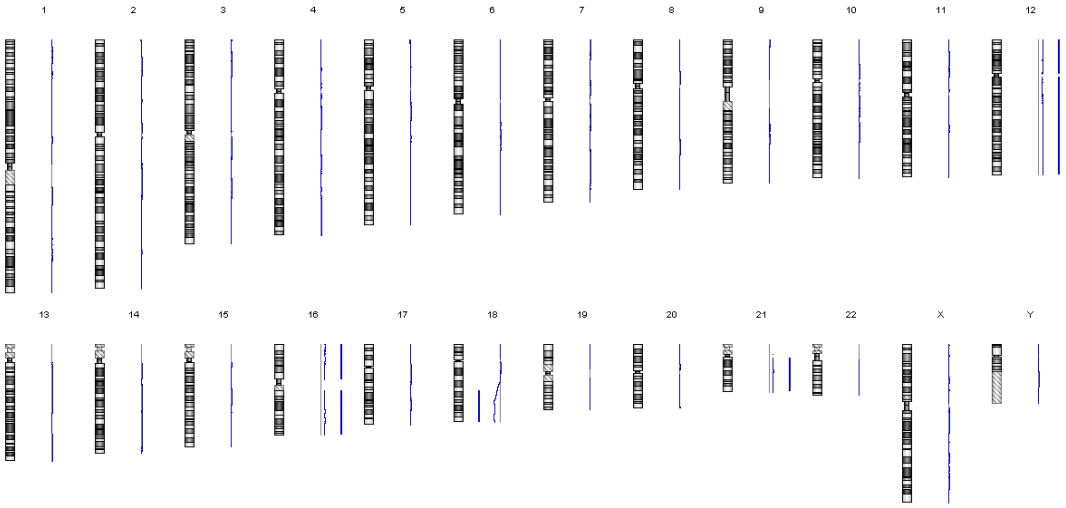
Matrix CGH profile of HD2026 liver metastasis tissue



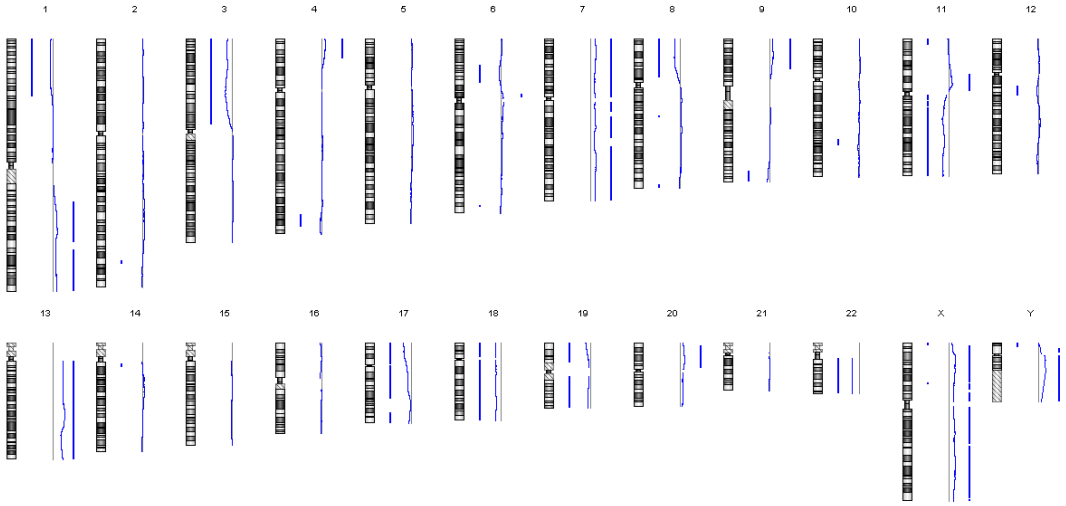
Matrix CGH profile of HD2095 liver metastasis tissue



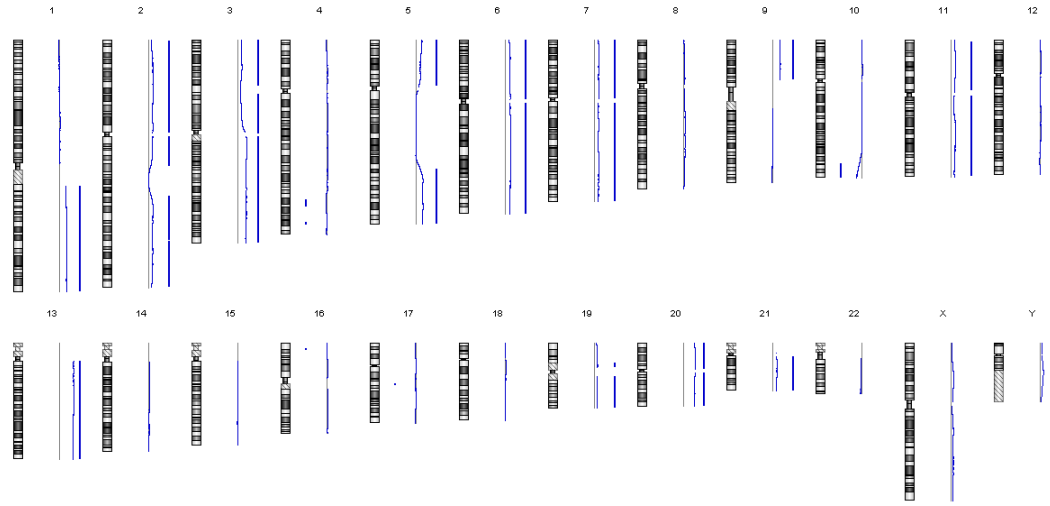
Matrix CGH profile of HD2126 liver metastasis tissue



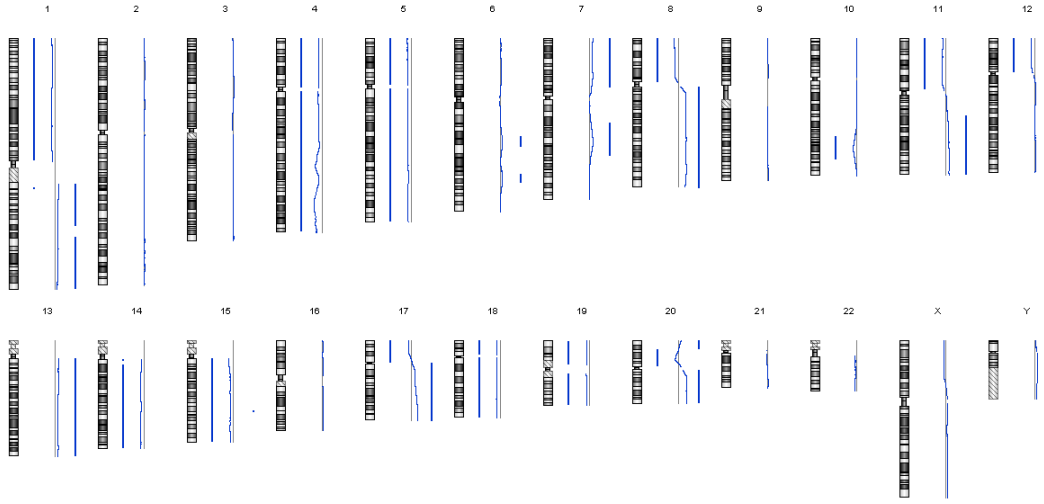
Matrix CGH profile of HD2224 tumor tissue



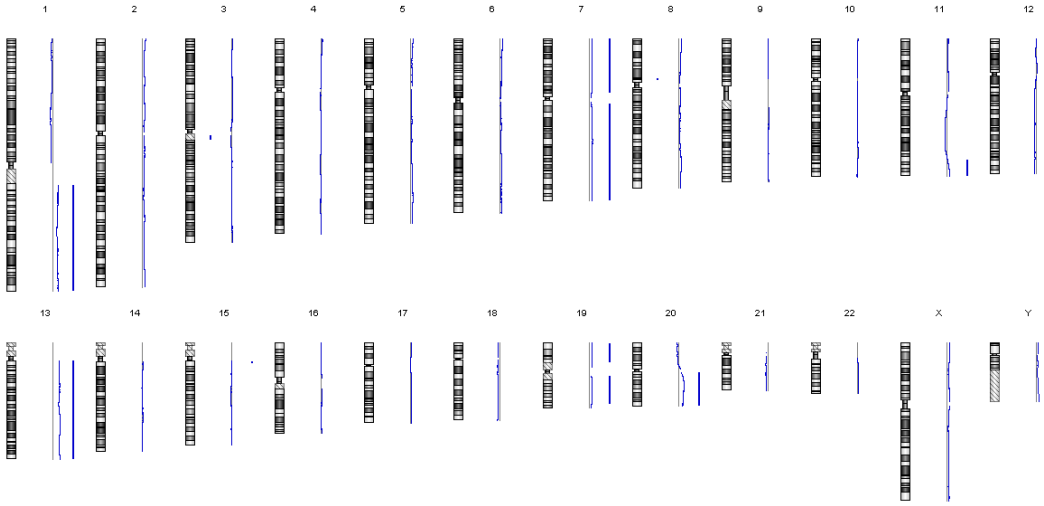
Matrix CGH profile of HD2231 liver metastasis tissue



Matrix CGH profile of HD2288 liver metastasis tissue



Matrix CGH profile of HD2328 liver metastasis tissue



Matrix CGH profile of HD2334 tumor tissue

Literature

1. Weitz, J. et al. Colorectal cancer. *Lancet* **365**, 153-165 (2005).
2. Boyle, P. & Ferlay, J. Cancer incidence and mortality in Europe, 2004. *Ann Oncol* **16**, 481-488 (2005).
3. Benson, A.B., 3rd Epidemiology, disease progression, and economic burden of colorectal cancer. *J Manag Care Pharm* **13**, S5-18 (2007).
4. Benson, A.B., 3rd et al. American Society of Clinical Oncology recommendations on adjuvant chemotherapy for stage II colon cancer. *J Clin Oncol* **22**, 3408-3419 (2004).
5. Chua, T.C., Liauw, W., Chu, F. & Morris, D.L. Viewing metastatic colorectal cancer as a curable chronic disease. *Am J Clin Oncol* **35**, 77-80 (2012).
6. Edge, S.B. & Compton, C.C. The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. *Ann Surg Oncol* **17**, 1471-1474 (2010).
7. Saif, M.W. & Chu, E. Biology of colorectal cancer. *Cancer J* **16**, 196-201 (2010).
8. Migliore, L., Migheli, F., Spisni, R. & Coppede, F. Genetics, cytogenetics, and epigenetics of colorectal cancer. *J Biomed Biotechnol* **2011**, 792362 (2011).
9. Sparks, A.B., Morin, P.J., Vogelstein, B. & Kinzler, K.W. Mutational analysis of the APC/beta-catenin/Tcf pathway in colorectal cancer. *Cancer Res* **58**, 1130-1134 (1998).
10. Fearon, E.R. Molecular genetics of colorectal cancer. *Annu Rev Pathol* **6**, 479-507 (2011).
11. Pino, M.S. & Chung, D.C. The chromosomal instability pathway in colon cancer. *Gastroenterology* **138**, 2059-2072 (2010).
12. Boland, C.R. & Goel, A. Microsatellite instability in colorectal cancer. *Gastroenterology* **138**, 2073-2087 e2073 (2010).
13. Markowitz, S. et al. Inactivation of the type II TGF-beta receptor in colon cancer cells with microsatellite instability. *Science* **268**, 1336-1338 (1995).
14. Massague, J., Blain, S.W. & Lo, R.S. TGFbeta signaling in growth control, cancer, and heritable disorders. *Cell* **103**, 295-309 (2000).
15. Markowitz, S.D. & Bertagnolli, M.M. Molecular origins of cancer: Molecular basis of colorectal cancer. *N Engl J Med* **361**, 2449-2460 (2009).
16. Roth, A.D. et al. Prognostic role of KRAS and BRAF in stage II and III resected colon cancer: results of the translational study on the PETACC-3, EORTC 40993, SAKK 60-00 trial. *J Clin Oncol* **28**, 466-474 (2010).
17. Baker, S.J. et al. p53 gene mutations occur in combination with 17p allelic deletions as late events in colorectal tumorigenesis. *Cancer Res* **50**, 7717-7722 (1990).
18. Greenblatt, M.S., Bennett, W.P., Hollstein, M. & Harris, C.C. Mutations in the p53 tumor suppressor gene: clues to cancer etiology and molecular pathogenesis. *Cancer Res* **54**, 4855-4878 (1994).
19. Samuels, Y. & Velculescu, V.E. Oncogenic mutations of PIK3CA in human cancers. *Cell Cycle* **3**, 1221-1224 (2004).
20. Michl, P. & Downward, J. Mechanisms of disease: PI3K/AKT signaling in gastrointestinal cancers. *Z Gastroenterol* **43**, 1133-1139 (2005).

21. Li, R. et al. Prognostic value of Akt-1 in human prostate cancer: a computerized quantitative assessment with quantum dot technology. *Clin Cancer Res* **15**, 3568-3573 (2009).
22. Tomita, Y. et al. Prognostic significance of activated AKT expression in soft-tissue sarcoma. *Clin Cancer Res* **12**, 3070-3077 (2006).
23. Dai, D.L., Martinka, M. & Li, G. Prognostic significance of activated Akt expression in melanoma: a clinicopathologic study of 292 cases. *J Clin Oncol* **23**, 1473-1482 (2005).
24. Baba, Y. et al. Phosphorylated AKT expression is associated with PIK3CA mutation, low stage, and favorable outcome in 717 colorectal cancers. *Cancer* **117**, 1399-1408 (2011).
25. Liu, M. & Chen, H. The role of microRNAs in colorectal cancer. *J Genet Genomics* **37**, 347-358 (2010).
26. Christiansen, J.J. & Rajasekaran, A.K. Reassessing epithelial to mesenchymal transition as a prerequisite for carcinoma invasion and metastasis. *Cancer Res* **66**, 8319-8326 (2006).
27. Tarin, D., Thompson, E.W. & Newgreen, D.F. The fallacy of epithelial mesenchymal transition in neoplasia. *Cancer Res* **65**, 5996-6000; discussion 6000-5991 (2005).
28. Polyak, K. & Weinberg, R.A. Transitions between epithelial and mesenchymal states: acquisition of malignant and stem cell traits. *Nat Rev Cancer* **9**, 265-273 (2009).
29. Huber, M.A., Kraut, N. & Beug, H. Molecular requirements for epithelial-mesenchymal transition during tumor progression. *Curr Opin Cell Biol* **17**, 548-558 (2005).
30. Boyer, B., Valles, A.M. & Edme, N. Induction and regulation of epithelial-mesenchymal transitions. *Biochem Pharmacol* **60**, 1091-1099 (2000).
31. Comijn, J. et al. The two-handed E box binding zinc finger protein SIP1 downregulates E-cadherin and induces invasion. *Mol Cell* **7**, 1267-1278 (2001).
32. Batlle, E. et al. The transcription factor snail is a repressor of E-cadherin gene expression in epithelial tumour cells. *Nat Cell Biol* **2**, 84-89 (2000).
33. Birchmeier, W. & Behrens, J. Cadherin expression in carcinomas: role in the formation of cell junctions and the prevention of invasiveness. *Biochim Biophys Acta* **1198**, 11-26 (1994).
34. Bohm, M., Totzeck, B., Birchmeier, W. & Wieland, I. Differences of E-cadherin expression levels and patterns in primary and metastatic human lung cancer. *Clin Exp Metastasis* **12**, 55-62 (1994).
35. Thiery, J.P. Epithelial-mesenchymal transitions in tumour progression. *Nat Rev Cancer* **2**, 442-454 (2002).
36. Peinado, H., Olmeda, D. & Cano, A. Snail, Zeb and bHLH factors in tumour progression: an alliance against the epithelial phenotype? *Nat Rev Cancer* **7**, 415-428 (2007).
37. Wu, Y. & Zhou, B.P. Snail: More than EMT. *Cell Adh Migr* **4**, 199-203 (2010).
38. Alves, C.C., Carneiro, F., Hoefler, H. & Becker, K.F. Role of the epithelial-mesenchymal transition regulator Slug in primary human cancers. *Front Biosci* **14**, 3035-3050 (2009).
39. Micalizzi, D.S., Farabaugh, S.M. & Ford, H.L. Epithelial-mesenchymal transition in cancer: parallels between normal development and tumor progression. *J Mammary Gland Biol Neoplasia* **15**, 117-134 (2010).
40. Yang, J. & Weinberg, R.A. Epithelial-mesenchymal transition: at the crossroads of development and tumor metastasis. *Dev Cell* **14**, 818-829 (2008).

41. Tarin, D. et al. Mechanisms of human tumor metastasis studied in patients with peritoneovenous shunts. *Cancer Res* **44**, 3584-3592 (1984).
42. Luzzi, K.J. et al. Multistep nature of metastatic inefficiency: dormancy of solitary cells after successful extravasation and limited survival of early micrometastases. *Am J Pathol* **153**, 865-873 (1998).
43. Brooks, D.E. The biorheology of tumor cells. *Biorheology* **21**, 85-91 (1984).
44. Taddei, M.L., Giannoni, E., Fiaschi, T. & Chiarugi, P. Anoikis: an emerging hallmark in health and diseases. *J Pathol* **226**, 380-393 (2012).
45. Hou, J.M. et al. Circulating tumor cells as a window on metastasis biology in lung cancer. *Am J Pathol* **178**, 989-996 (2011).
46. Andratschke, M., Pauli, C., Stein, M., Chaubal, S. & Wollenberg, B. MHC-class I antigen expression on micrometastases in bone marrow of patients with head and neck squamous cell cancer. *Anticancer Res* **23**, 1467-1471 (2003).
47. Nieswandt, B., Hafner, M., Echtenacher, B. & Mannel, D.N. Lysis of tumor cells by natural killer cells in mice is impeded by platelets. *Cancer Res* **59**, 1295-1300 (1999).
48. Placke, T. et al. Platelet-derived MHC class I confers a pseudonormal phenotype to cancer cells that subverts the antitumor reactivity of natural killer immune cells. *Cancer Res* **72**, 440-448 (2012).
49. Pawelek, J.M. & Chakraborty, A.K. The cancer cell--leukocyte fusion theory of metastasis. *Adv Cancer Res* **101**, 397-444 (2008).
50. Colombano, S.P. & Reese, P.A. The cascade theory of metastatic spread: are there generalizing sites? *Cancer* **46**, 2312-2314 (1980).
51. Makrilia, N., Kollias, A., Manolopoulos, L. & Syrigos, K. Cell adhesion molecules: role and clinical significance in cancer. *Cancer Invest* **27**, 1023-1037 (2009).
52. Muller, A. et al. Involvement of chemokine receptors in breast cancer metastasis. *Nature* **410**, 50-56 (2001).
53. Sastre, J. et al. Circulating tumor cells in colorectal cancer: correlation with clinical and pathological variables. *Ann Oncol* **19**, 935-938 (2008).
54. Fidler, I.J. The pathogenesis of cancer metastasis: the 'seed and soil' hypothesis revisited. *Nat Rev Cancer* **3**, 453-458 (2003).
55. Psaila, B., Kaplan, R.N., Port, E.R. & Lyden, D. Priming the 'soil' for breast cancer metastasis: the pre-metastatic niche. *Breast Dis* **26**, 65-74 (2006).
56. Kaplan, R.N. et al. VEGFR1-positive haematopoietic bone marrow progenitors initiate the pre-metastatic niche. *Nature* **438**, 820-827 (2005).
57. Tsuji, T. et al. Epithelial-mesenchymal transition induced by growth suppressor p12CDK2-AP1 promotes tumor cell local invasion but suppresses distant colony growth. *Cancer Res* **68**, 10377-10386 (2008).
58. Galandiuk, S. et al. Patterns of recurrence after curative resection of carcinoma of the colon and rectum. *Surg Gynecol Obstet* **174**, 27-32 (1992).
59. Li, J. et al. Tumor heterogeneity in neoplasms of breast, colon, and skin. *BMC Res Notes* **3**, 321 (2010).
60. Grady, W.M. & Markowitz, S. Genomic instability and colorectal cancer. *Curr Opin Gastroenterol* **16**, 62-67 (2000).
61. Visvader, J.E. & Lindeman, G.J. Cancer stem cells in solid tumours: accumulating evidence and unresolved questions. *Nat Rev Cancer* **8**, 755-768 (2008).
62. Pang, R. et al. A subpopulation of CD26+ cancer stem cells with metastatic capacity in human colorectal cancer. *Cell Stem Cell* **6**, 603-615 (2010).

63. Ricci-Vitiani, L. et al. Identification and expansion of human colon-cancer-initiating cells. *Nature* **445**, 111-115 (2007).
64. Kelly, P.N., Dakic, A., Adams, J.M., Nutt, S.L. & Strasser, A. Tumor growth need not be driven by rare cancer stem cells. *Science* **317**, 337 (2007).
65. Hill, R.P. Identifying cancer stem cells in solid tumors: case not proven. *Cancer Res* **66**, 1891-1895; discussion 1890 (2006).
66. Quintana, E. et al. Efficient tumour formation by single human melanoma cells. *Nature* **456**, 593-598 (2008).
67. Brabletz, T., Jung, A., Spaderna, S., Hlubek, F. & Kirchner, T. Opinion: migrating cancer stem cells - an integrated concept of malignant tumour progression. *Nat Rev Cancer* **5**, 744-749 (2005).
68. Ovaska, J., Jarvinen, H., Kujari, H., Perttila, I. & Mecklin, J.P. Follow-up of patients operated on for colorectal carcinoma. *Am J Surg* **159**, 593-596 (1990).
69. Wolpin, B.M. & Mayer, R.J. Systemic treatment of colorectal cancer. *Gastroenterology* **134**, 1296-1310 (2008).
70. Klein, C.A. Parallel progression of primary tumours and metastases. *Nat Rev Cancer* **9**, 302-312 (2009).
71. Lianidou, E.S. & Markou, A. Circulating Tumor Cells in Breast Cancer: Detection Systems, Molecular Characterization, and Future Challenges. *Clin Chem* (2011).
72. Yu, M., Stott, S., Toner, M., Maheswaran, S. & Haber, D.A. Circulating tumor cells: approaches to isolation and characterization. *J Cell Biol* **192**, 373-382 (2011).
73. Cristofanilli, M. et al. Circulating tumor cells, disease progression, and survival in metastatic breast cancer. *N Engl J Med* **351**, 781-791 (2004).
74. Sequist, L.V., Nagrath, S., Toner, M., Haber, D.A. & Lynch, T.J. The CTC-chip: an exciting new tool to detect circulating tumor cells in lung cancer patients. *J Thorac Oncol* **4**, 281-283 (2009).
75. Talasaz, A.H. et al. Isolating highly enriched populations of circulating epithelial cells and other rare cells from blood using a magnetic sweeper device. *Proc Natl Acad Sci U S A* **106**, 3970-3975 (2009).
76. Vona, G. et al. Isolation by size of epithelial tumor cells : a new method for the immunomorphological and molecular characterization of circulatingtumor cells. *Am J Pathol* **156**, 57-63 (2000).
77. Marrinucci, D. et al. Case study of the morphologic variation of circulating tumor cells. *Hum Pathol* **38**, 514-519 (2007).
78. Pantel, K. et al. Circulating Epithelial Cells in Patients with Benign Colon Diseases. *Clin Chem* (2011).
79. Schmitt, M. & Foekens, J.A. Circulating tumor cells in blood of primary breast cancer patients assessed by a novel RT-PCR test kit and comparison with status of bone marrow-disseminated tumor cells. *Breast Cancer Res* **11**, 109 (2009).
80. Shimomura, M. et al. Adequate lymph node examination is essential to ensure the prognostic value of the lymph node ratio in patients with stage III colorectal cancer. *Surg Today* **41**, 1370-1379 (2011).
81. Weitz, J. et al. Detection of disseminated colorectal cancer cells in lymph nodes, blood and bone marrow. *Clin Cancer Res* **5**, 1830-1836 (1999).
82. Wolfrum, F., Vogel, I., Fandrich, F. & Kalthoff, H. Detection and clinical implications of minimal residual disease in gastro-intestinal cancer. *Langenbecks Arch Surg* **390**, 430-441 (2005).

83. Nicastri, D.G., Doucette, J.T., Godfrey, T.E. & Hughes, S.J. Is occult lymph node disease in colorectal cancer patients clinically significant? A review of the relevant literature. *J Mol Diagn* **9**, 563-571 (2007).
84. Rahbari, N.N. et al. Meta-analysis shows that detection of circulating tumor cells indicates poor prognosis in patients with colorectal cancer. *Gastroenterology* **138**, 1714-1726 (2010).
85. Allard, W.J. et al. 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**, 6897-6904 (2004).
86. Cohen, S.J. et al. Prognostic significance of circulating tumor cells in patients with metastatic colorectal cancer. *Ann Oncol* **20**, 1223-1229 (2009).
87. Miller, M.C., Doyle, G.V. & Terstappen, L.W. Significance of Circulating Tumor Cells Detected by the CellSearch System in Patients with Metastatic Breast Colorectal and Prostate Cancer. *J Oncol* **2010**, 617421 (2010).
88. Chambers, A.F., Groom, A.C. & MacDonald, I.C. Dissemination and growth of cancer cells in metastatic sites. *Nat Rev Cancer* **2**, 563-572 (2002).
89. Peach, G., Kim, C., Zacharakis, E., Purkayastha, S. & Ziprin, P. Prognostic significance of circulating tumour cells following surgical resection of colorectal cancers: a systematic review. *Br J Cancer* **102**, 1327-1334 (2010).
90. Riethdorf, S., Wikman, H. & Pantel, K. Review: Biological relevance of disseminated tumor cells in cancer patients. *Int J Cancer* **123**, 1991-2006 (2008).
91. Pantel, K. et al. Detection and clinical implications of early systemic tumor cell dissemination in breast cancer. *Clin Cancer Res* **9**, 6326-6334 (2003).
92. Zia, A., Schildberg, F.W. & Funke, I. MHC class I negative phenotype of disseminated tumor cells in bone marrow is associated with poor survival in R0M0 breast cancer patients. *Int J Cancer* **93**, 566-570 (2001).
93. Wulfing, P. et al. HER2-positive circulating tumor cells indicate poor clinical outcome in stage I to III breast cancer patients. *Clin Cancer Res* **12**, 1715-1720 (2006).
94. Pierga, J.Y. et al. Real-time quantitative PCR determination of urokinase-type plasminogen activator receptor (uPAR) expression of isolated micrometastatic cells from bone marrow of breast cancer patients. *Int J Cancer* **114**, 291-298 (2005).
95. Theodoropoulos, P.A. et al. Circulating tumor cells with a putative stem cell phenotype in peripheral blood of patients with breast cancer. *Cancer Lett* **288**, 99-106 (2010).
96. Muller, P., Weckermann, D., Riethmuller, G. & Schlimok, G. Detection of genetic alterations in micrometastatic cells in bone marrow of cancer patients by fluorescence in situ hybridization. *Cancer Genet Cytogenet* **88**, 8-16 (1996).
97. Klein, C.A. et al. Genetic heterogeneity of single disseminated tumour cells in minimal residual cancer. *Lancet* **360**, 683-689 (2002).
98. Al-Mulla, F. et al. Heterogeneity of mutant versus wild-type Ki-ras in primary and metastatic colorectal carcinomas, and association of codon-12 valine with early mortality. *J Pathol* **185**, 130-138 (1998).
99. Stoecklein, N.H. & Klein, C.A. Genetic disparity between primary tumours, disseminated tumour cells, and manifest metastasis. *Int J Cancer* **126**, 589-598 (2010).
100. Smirnov, D.A. et al. Global gene expression profiling of circulating tumor cells. *Cancer Res* **65**, 4993-4997 (2005).

101. Sieuwerts, A.M. et al. mRNA and microRNA expression profiles in circulating tumor cells and primary tumors of metastatic breast cancer patients. *Clin Cancer Res* **17**, 3600-3618 (2011).
102. Watson, M.A. et al. Isolation and molecular profiling of bone marrow micrometastases identifies TWIST1 as a marker of early tumor relapse in breast cancer patients. *Clin Cancer Res* **13**, 5001-5009 (2007).
103. Klein, C.A. et al. Combined transcriptome and genome analysis of single micrometastatic cells. *Nat Biotechnol* **20**, 387-392 (2002).
104. Rahbari, N.N. et al. Compartmental Differences of Circulating Tumor Cells in Colorectal Cancer. *Ann Surg Oncol* (2012).
105. Kim, M.J. et al. Association of CD47 with natural killer cell-mediated cytotoxicity of head-and-neck squamous cell carcinoma lines. *Tumour Biol* **29**, 28-34 (2008).
106. Losi, L., Baisse, B., Bouzourene, H. & Benhattar, J. Evolution of intratumoral genetic heterogeneity during colorectal cancer progression. *Carcinogenesis* **26**, 916-922 (2005).
107. Heng, H.H. et al. Genetic and epigenetic heterogeneity in cancer: the ultimate challenge for drug therapy. *Curr Drug Targets* **11**, 1304-1316 (2010).
108. Klein, C.A. et al. Comparative genomic hybridization, loss of heterozygosity, and DNA sequence analysis of single cells. *Proc Natl Acad Sci U S A* **96**, 4494-4499 (1999).
109. Sargent, D.J. et al. Defective mismatch repair as a predictive marker for lack of efficacy of fluorouracil-based adjuvant therapy in colon cancer. *J Clin Oncol* **28**, 3219-3226 (2010).
110. Oki, E. et al. Chemosensitivity and survival in gastric cancer patients with microsatellite instability. *Ann Surg Oncol* **16**, 2510-2515 (2009).
111. Umar, A. et al. Revised Bethesda Guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. *J Natl Cancer Inst* **96**, 261-268 (2004).
112. Shemirani, A.I. et al. Simplified MSI Marker Panel for Diagnosis of Colorectal Cancer. *Asian Pac J Cancer Prev* **12**, 2101-2104 (2011).
113. Woerner, S.M. et al. Microsatellite instability of selective target genes in HNPCC-associated colon adenomas. *Oncogene* **24**, 2525-2535 (2005).
114. Halford, S.E. et al. MSI-low, a real phenomenon which varies in frequency among cancer types. *J Pathol* **201**, 389-394 (2003).
115. Halford, S. et al. Low-level microsatellite instability occurs in most colorectal cancers and is a nonrandomly distributed quantitative trait. *Cancer Res* **62**, 53-57 (2002).
116. Tanaka, T. et al. Chromosome 18q deletion as a novel molecular predictor for colorectal cancer with simultaneous hepatic metastasis. *Diagn Mol Pathol* **18**, 219-225 (2009).
117. Knosel, T. et al. Genetic imbalances with impact on survival in colorectal cancer patients. *Histopathology* **43**, 323-331 (2003).
118. Catenacci, D.V., Kozloff, M., Kindler, H.L. & Polite, B. Personalized colon cancer care in 2010. *Semin Oncol* **38**, 284-308 (2011).
119. Naomoto, Y., Kondo, H., Tanaka, N. & Orita, K. Novel experimental models of human cancer metastasis in nude mice: lung metastasis, intraabdominal carcinomatosis with ascites, and liver metastasis. *J Cancer Res Clin Oncol* **113**, 544-549 (1987).

120. Thalheimer, A. et al. The intraportal injection model: a practical animal model for hepatic metastases and tumor cell dissemination in human colon cancer. *BMC Cancer* **9**, 29 (2009).
121. Flatmark, K., Maelandsmo, G.M., Martinsen, M., Rasmussen, H. & Fodstad, O. Twelve colorectal cancer cell lines exhibit highly variable growth and metastatic capacities in an orthotopic model in nude mice. *Eur J Cancer* **40**, 1593-1598 (2004).
122. Unemori, E.N., Ways, N. & Pitelka, D.R. Metastasis of murine mammary tumour lines from the mammary gland and ectopic sites. *Br J Cancer* **49**, 603-614 (1984).
123. Kubota, T. Metastatic models of human cancer xenografted in the nude mouse: the importance of orthotopic transplantation. *J Cell Biochem* **56**, 4-8 (1994).
124. Glinskii, A.B. et al. Viable circulating metastatic cells produced in orthotopic but not ectopic prostate cancer models. *Cancer Res* **63**, 4239-4243 (2003).
125. Pantelouris, E.M. Observations on the immunobiology of 'nude' mice. *Immunology* **20**, 247-252 (1971).
126. Cespedes, M.V. et al. Orthotopic microinjection of human colon cancer cells in nude mice induces tumor foci in all clinically relevant metastatic sites. *Am J Pathol* **170**, 1077-1085 (2007).
127. Sasaki, H. et al. Orthotopic implantation mouse model and cDNA microarray analysis indicates several genes potentially involved in lymph node metastasis of colorectal cancer. *Cancer Sci* **99**, 711-719 (2008).
128. Shultz, L.D., Ishikawa, F. & Greiner, D.L. Humanized mice in translational biomedical research. *Nat Rev Immunol* **7**, 118-130 (2007).
129. Morikawa, K., Walker, S.M., Jessup, J.M. & Fidler, I.J. In vivo selection of highly metastatic cells from surgical specimens of different primary human colon carcinomas implanted into nude mice. *Cancer Res* **48**, 1943-1948 (1988).
130. Kuhn, S. et al. A complex of EpCAM, claudin-7, CD44 variant isoforms, and tetraspanins promotes colorectal cancer progression. *Mol Cancer Res* **5**, 553-567 (2007).
131. Koch, M. et al. Comparative analysis of tumor cell dissemination in mesenteric, central, and peripheral venous blood in patients with colorectal cancer. *Arch Surg* **136**, 85-89 (2001).
132. Taniguchi, T., Makino, M., Suzuki, K. & Kaibara, N. Prognostic significance of reverse transcriptase-polymerase chain reaction measurement of carcinoembryonic antigen mRNA levels in tumor drainage blood and peripheral blood of patients with colorectal carcinoma. *Cancer* **89**, 970-976 (2000).
133. Fujita, S., Kudo, N., Akasu, T. & Moriya, Y. Detection of cytokeratin 19 and 20 mRNA in peripheral and mesenteric blood from colorectal cancer patients and their prognosis. *Int J Colorectal Dis* **16**, 141-146 (2001).
134. Iinuma, H. et al. Usefulness and clinical significance of quantitative real-time RT-PCR to detect isolated tumor cells in the peripheral blood and tumor drainage blood of patients with colorectal cancer. *Int J Oncol* **28**, 297-306 (2006).
135. Sato, H. et al. Significance of serum concentrations of E-selectin and CA19-9 in the prognosis of colorectal cancer. *Jpn J Clin Oncol* **40**, 1073-1080 (2010).
136. Dabelsteen, E. Cell surface carbohydrates as prognostic markers in human carcinomas. *J Pathol* **179**, 358-369 (1996).
137. Park, I.J., Choi, G.S. & Jun, S.H. Prognostic value of serum tumor antigen CA19-9 after curative resection of colorectal cancer. *Anticancer Res* **29**, 4303-4308 (2009).

138. Sieuwerts, A.M. et al. Anti-epithelial cell adhesion molecule antibodies and the detection of circulating normal-like breast tumor cells. *J Natl Cancer Inst* **101**, 61-66 (2009).
139. Krebs, M.G. et al. Analysis of Circulating Tumor Cells in Patients with Non-small Cell Lung Cancer Using Epithelial Marker-Dependent and -Independent Approaches. *J Thorac Oncol* (2011).
140. van de Stolpe, A., Pantel, K., Sleijfer, S., Terstappen, L.W. & den Toonder, J.M. Circulating tumor cell isolation and diagnostics: toward routine clinical use. *Cancer Res* **71**, 5955-5960 (2011).
141. Raj, A. & van Oudenaarden, A. Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell* **135**, 216-226 (2008).
142. Elzagheid, A. et al. E-cadherin expression pattern in primary colorectal carcinomas and their metastases reflects disease outcome. *World J Gastroenterol* **12**, 4304-4309 (2006).
143. Tsuji, T., Ibaragi, S. & Hu, G.F. Epithelial-mesenchymal transition and cell cooperativity in metastasis. *Cancer Res* **69**, 7135-7139 (2009).
144. Mehes, G., Witt, A., Kubista, E. & Ambros, P.F. Circulating breast cancer cells are frequently apoptotic. *Am J Pathol* **159**, 17-20 (2001).
145. Nubel, T. et al. Claudin-7 regulates EpCAM-mediated functions in tumor progression. *Mol Cancer Res* **7**, 285-299 (2009).
146. Muller, V. et al. Circulating tumor cells in breast cancer: correlation to bone marrow micrometastases, heterogeneous response to systemic therapy and low proliferative activity. *Clin Cancer Res* **11**, 3678-3685 (2005).
147. Fuchs, S.Y., Ougolkov, A.V., Spiegelman, V.S. & Minamoto, T. Oncogenic beta-catenin signaling networks in colorectal cancer. *Cell Cycle* **4**, 1522-1539 (2005).
148. Kikuchi, A., Kishida, S. & Yamamoto, H. Regulation of Wnt signaling by protein-protein interaction and post-translational modifications. *Exp Mol Med* **38**, 1-10 (2006).
149. Pantel, K. et al. Differential expression of proliferation-associated molecules in individual micrometastatic carcinoma cells. *J Natl Cancer Inst* **85**, 1419-1424 (1993).
150. Solakoglu, O. et al. Heterogeneous proliferative potential of occult metastatic cells in bone marrow of patients with solid epithelial tumors. *Proc Natl Acad Sci U S A* **99**, 2246-2251 (2002).
151. Di Fiore, F. et al. Clinical relevance of KRAS mutation detection in metastatic colorectal cancer treated by Cetuximab plus chemotherapy. *Br J Cancer* **96**, 1166-1169 (2007).
152. Lievre, A. et al. KRAS mutation status is predictive of response to cetuximab therapy in colorectal cancer. *Cancer Res* **66**, 3992-3995 (2006).
153. Di Nicolantonio, F. et al. Wild-type BRAF is required for response to panitumumab or cetuximab in metastatic colorectal cancer. *J Clin Oncol* **26**, 5705-5712 (2008).
154. Siena, S., Sartore-Bianchi, A., Di Nicolantonio, F., Balfour, J. & Bardelli, A. Biomarkers predicting clinical outcome of epidermal growth factor receptor-targeted therapy in metastatic colorectal cancer. *J Natl Cancer Inst* **101**, 1308-1324 (2009).
155. Vidaurreta, M. et al. Colorectal carcinoma prognosis can be predicted by alterations in gene p53 exons 5 and 8. *Int J Colorectal Dis* **23**, 581-586 (2008).

156. Mollevi, D.G. et al. Mutations in TP53 are a prognostic factor in colorectal hepatic metastases undergoing surgical resection. *Carcinogenesis* **28**, 1241-1246 (2007).
157. Elsaleh, H. et al. P53 alteration and microsatellite instability have predictive value for survival benefit from chemotherapy in stage III colorectal carcinoma. *Clin Cancer Res* **7**, 1343-1349 (2001).
158. Samowitz, W.S., Curtin, K., Neuhausen, S., Schaffer, D. & Slattery, M.L. Prognostic implications of BAX and TGFBRII mutations in colon cancers with microsatellite instability. *Genes Chromosomes Cancer* **35**, 368-371 (2002).
159. Benatti, P. et al. Microsatellite instability and colorectal cancer prognosis. *Clin Cancer Res* **11**, 8332-8340 (2005).
160. Brosens, R.P. et al. Deletion of chromosome 4q predicts outcome in stage II colon cancer patients. *Cell Oncol (Dordr)* **34**, 215-223 (2011).
161. Watanabe, T. et al. Molecular predictors of survival after adjuvant chemotherapy for colon cancer. *N Engl J Med* **344**, 1196-1206 (2001).
162. Chen, Z. et al. Chromosomal copy number alterations are associated with tumor response to chemoradiation in locally advanced rectal cancer. *Genes Chromosomes Cancer* **50**, 689-699 (2011).
163. Fehm, T. et al. Cytogenetic evidence that circulating epithelial cells in patients with carcinoma are malignant. *Clin Cancer Res* **8**, 2073-2084 (2002).
164. Baker, S.J. et al. Chromosome 17 deletions and p53 gene mutations in colorectal carcinomas. *Science* **244**, 217-221 (1989).
165. Vogelstein, B. & Kinzler, K.W. Cancer genes and the pathways they control. *Nat Med* **10**, 789-799 (2004).
166. Schardt, J.A. et al. Genomic analysis of single cytokeratin-positive cells from bone marrow reveals early mutational events in breast cancer. *Cancer Cell* **8**, 227-239 (2005).
167. Albanese, I. et al. Heterogeneity within and between primary colorectal carcinomas and matched metastases as revealed by analysis of Ki-ras and p53 mutations. *Biochem Biophys Res Commun* **325**, 784-791 (2004).
168. Artale, S. et al. Mutations of KRAS and BRAF in primary and matched metastatic sites of colorectal cancer. *J Clin Oncol* **26**, 4217-4219 (2008).
169. Santini, D. et al. High concordance of KRAS status between primary colorectal tumors and related metastatic sites: implications for clinical practice. *Oncologist* **13**, 1270-1275 (2008).
170. Baldus, S.E. et al. Prevalence and heterogeneity of KRAS, BRAF, and PIK3CA mutations in primary colorectal adenocarcinomas and their corresponding metastases. *Clin Cancer Res* **16**, 790-799 (2010).
171. Yao, J., Goh, H.S. & Smith, D.R. p53 mutations in primary colorectal adenocarcinomas and liver metastases. *Br J Surg* **83**, 1245-1246 (1996).
172. Zhang, J.S., Caplin, S., Bosman, F.T. & Benhattar, J. Genetic diversity at the p53 locus between primary human colorectal adenocarcinomas and their lymph-node metastases. *Int J Cancer* **70**, 674-678 (1997).
173. Schmidt-Kittler, O. et al. From latent disseminated cells to overt metastasis: genetic analysis of systemic breast cancer progression. *Proc Natl Acad Sci U S A* **100**, 7737-7742 (2003).
174. Kraus, J., Pantel, K., Pinkel, D., Albertson, D.G. & Speicher, M.R. High-resolution genomic profiling of occult micrometastatic tumor cells. *Genes Chromosomes Cancer* **36**, 159-166 (2003).

175. Gangnus, R., Langer, S., Breit, E., Pantel, K. & Speicher, M.R. Genomic profiling of viable and proliferative micrometastatic cells from early-stage breast cancer patients. *Clin Cancer Res* **10**, 3457-3464 (2004).
176. Gorelik, E., Gunji, Y. & Goldfarb, R.H. Interaction of tumor cells and immune system in the metastatic process. *Biochem Cell Biol* **66**, 617-625 (1988).
177. Khan, Z.A. et al. P53 mutations in primary and metastatic tumors and circulating tumor cells from colorectal carcinoma patients. *Clin Cancer Res* **6**, 3499-3504 (2000).
178. Meng, S. et al. HER-2 gene amplification can be acquired as breast cancer progresses. *Proc Natl Acad Sci U S A* **101**, 9393-9398 (2004).
179. Hayes, D.F. et al. Monitoring expression of HER-2 on circulating epithelial cells in patients with advanced breast cancer. *Int J Oncol* **21**, 1111-1117 (2002).
180. Eliane, J.P. et al. Monitoring serial changes in circulating human breast cancer cells in murine xenograft models. *Cancer Res* **68**, 5529-5532 (2008).
181. Helzer, K.T. et al. Circulating tumor cells are transcriptionally similar to the primary tumor in a murine prostate model. *Cancer Res* **69**, 7860-7866 (2009).
182. Howard, E.W. et al. Decreased adhesiveness, resistance to anoikis and suppression of GRP94 are integral to the survival of circulating tumor cells in prostate cancer. *Clin Exp Metastasis* **25**, 497-508 (2008).
183. Wang, D. et al. Downregulation of E-Cadherin enhances proliferation of head and neck cancer through transcriptional regulation of EGFR. *Mol Cancer* **10**, 116 (2011).
184. Misra, S. et al. Hyaluronan-CD44 interactions as potential targets for cancer therapy. *FEBS J* **278**, 1429-1443 (2011).
185. Boland, C.R. et al. A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. *Cancer Res* **58**, 5248-5257 (1998).
186. Findeisen, P. et al. T25 repeat in the 3' untranslated region of the CASP2 gene: a sensitive and specific marker for microsatellite instability in colorectal cancer. *Cancer Res* **65**, 8072-8078 (2005).

Acknowledgment

I am very grateful to everyone who supported my work on this thesis. In particular I would like to thank Prof. Dr. Jürgen Weitz for giving me the chance to work on a very challenging field of research in his lab. Furthermore, I have to thank Prof. Dr. Philipp Beckhove for his supervision.

Thank you Prof. Dr. Moritz Koch for your help and advises.

Thank you, Dr. Gunnar Steinert for your help with the project as well as for reading the manuscript and your suggestions for improvement.

Thanks to the whole lab team for the pleasant working atmosphere. I really learned a lot about playing Schnauz.

Thanks to the external collaboration partners Dr. Matthias Kloor and Anita Voigt as well as Prof. Dr. N. Stöcklein and Bianca Behrens for their contributions and support.

Furthermore, I will not omit to mention Martin Mollenhauer and Markus Stauch who became good friends and helped me to overcome some times of disappointment and despair.

Finally, I am very grateful to my whole family and in particular to my girl friend Janine Kästner for her support and comfort during a hard time.

Declaration

I have written this thesis independently, solely based on the literature, methods and devices mentioned in the chapters and the appendix.