

Dissertation

**Establishment of new methods for
detection and isolation of circulating tumor
cells and circulating cell-free tumor DNA in
cancer patients**

submitted by

Mag.rer.nat

Eva Maria Hoffmann

for the Academic Degree of

Doctor of Philosophy

(Ph.D.)

at the

Medical University of Graz

Institute of Human Genetics

under supervision of

Univ.-Prof. Dr. med. Michael Speicher

2014

Declaration

I hereby declare that this thesis is my own original work and that I have fully acknowledged by name all individuals and organizations that have contributed to the research for this thesis. Due acknowledgement has been made in the text to all other material used. Throughout this thesis and in all related publications I have followed the guidelines of “Good Scientific Practice”.

Graz,

*Der Beginn aller Wissenschaften ist das Erstaunen, dass die
Dinge sind, wie sie sind. (Aristoteles)*

Meinem Großvater Franz Eckhart gewidmet.

ACKNOWLEDGEMENTS

Zu Beginn möchte ich Herrn Prof. Speicher danken. Für seine Unterstützung, das übermittelte Wissen und seine immer offene Tür. Für sein Vertrauen und den Glauben an mich und daran, dass diese Zellen irgendwann leuchten werden. Vielen Dank für Ihre Geduld!

Außerdem möchte ich Frau Prof. Dagmar Kratky und Herrn Prof. Jochen Geigl, den Mitgliedern meines Thesis committees, danken. Danke Jochen, für die unzähligen wissenschaftlichen Diskussionen und deinen freundschaftlichen Rat, wenn es darauf angekommen ist.

Ich bedanke mich bei Ellen und Peter für ihre Hilfsbereitschaft in all den aussichtslos erscheinenden Momenten.

Vielen, vielen Dank an Martina und Julie. Danke, dass mein Bauch vor Lachen schmerzen durfte. Danke, dass ihr eure Ohren und euer Herz für mich geöffnet habt. Danke, dass ihr diesen Weg mit mir gemeinsam gegangen seid.

Danke auch an Christine, Sam, Sumitra, Jelena, Maryam, Ricarda und Kristina, für die kuschelige Atmosphäre in unserem Büro.

Ein großes Dankeschön gebührt auch Leni, Anja und Julia, die immer da waren, wenn ich über weltliche Dinge reden musste.

Und Wolfgang – dir danke ich am allermeisten. Du warst mein geduldiges Übungspublikum, als ich mich für Thesis committee meetings, Doctoral days und Journal Clubs vorbereiten musste. Du hast dich mit mir gefreut und mit mir gelitten und auch wenn es mir gelungen ist diese Dissertation zu schreiben, so fehlen mir die Worte um auszudrücken, wie sehr ich dir für alles danke.

Zu guter Letzt danke ich meinen Eltern, die immer hinter mir standen, aber natürlich auch den tollsten Schwestern und den süßesten Nichten dieser Welt. Dankeschön!

Index

1. Abstract	I
1.1. Abstract	I
1.2. Zusammenfassung	II
1.3. Abbreviations	IV
Establishment of tumor-specific copy number alterations from plasma DNA of patients with cancer	2
2. GENERAL INTRODUCTION	3
2.1. Colorectal Cancer	3
2.1.1. Epidemiology of colorectal cancer	3
2.1.2. Stages of colorectal cancer	4
2.3. Cancer Metastasis	5
2.4. Circulating DNA	8
2.4.1. Circulating tumor DNA (ctDNA) in cancer patients	8
2.4.2. Circulating tumor DNA	9
2.4.3. The role of ctDNA as biomarker in carcinogenesis	11
2.5. Circulating tumor cells (CTCs)	12
3. AIM OF THE THESIS	17
3.1. Specific aims:	17
3.1.1. The following research questions were defined:	17
4. Material and Methods	18
4.1. Material	18
4.1.1. Chemicals and Media	18
4.1.2. Kits and Standards	19
4.1.3. Bacterial strains	21
4.1.4. Plasmids	21
4.1.5. Cell lines	21
4.1.6. Oligonucleotides	22
4.1.7. Laboratory equipment	23
4.1.8. Buffers	25
4.1.9. Software, Database	25
4.2. Methods	26

4.2.1.	<i>Sample collection</i>	26
4.2.2.	<i>Plasma preparation</i>	27
4.2.3.	<i>Isolation of cell-free circulating DNA from plasma</i>	28
4.2.4.	<i>Plasma-DNA analysis</i>	29
4.2.5.	<i>DNA Isolation Methods</i>	36
4.2.6.	<i>DNA Amplification Methods</i>	36
4.2.7.	<i>Detection of CTCs and whole genome amplification</i>	37
4.2.8.	<i>DNA Purification Methods</i>	38
4.2.9.	<i>. DNA Labeling Methods</i>	39
4.2.10.	<i>Comparative Genomic Hybridization, CGH</i>	41
4.2.11.	<i>BEAMing (Beads, Emulsions, Amplification, and Magnetics) Assay</i>	45
4.2.12.	<i>Deep sequencing with the 454 GS FLX genome sequencer</i>	45
5.	Results	47
5.1.1.	<i>Analysis of patients after sex-mismatched bone marrow transplantation</i>	47
5.1.2.	<i>Analysis of patients with CRC</i>	49
6.	Discussion	61
7.	DEVELOPMENT OF AN ADENOVIRAL VECTOR FOR DETECTING TELOMERASE POSITIVE CIRCULATING TUMOR CELLS	65
7.1.	<i>Telomere structure and function</i>	65
7.2.	<i>Telomerase</i>	67
7.3.	<i>Telomerase activity</i>	69
7.4.	<i>Telomerase promoter</i>	70
7.5.	<i>Adenovirus</i>	71
7.5.1.	<i>Adenoviral structure</i>	72
7.5.2.	<i>Adenoviral genome</i>	72
7.5.3.	<i>Construction of a telomerase-specific vector</i>	74
7.5.4.	<i>Detection of circulating tumor cells using telomerase-specific adenovirus</i>	74
8.	AIMS OF THE PROJECT	77
8.1.	<i>General aims</i>	77
8.2.	<i>Specific aims</i>	77
9.	Methods	78
9.1.	<i>Bacteria work</i>	78
9.1.1.	<i>Overnight culture</i>	78

9.2.	DNA purification methods.....	78
9.2.1.	GenElute® Plasmid Miniprep Kit (Sigma Aldrich).....	78
9.2.2.	GenElute® PCR Clean Up Kit.....	79
9.2.3.	Purification of DNA extracted from agarose gel.....	79
9.3.	DNA amplification methods.....	80
9.3.1.	Polymerase Chain reaction (PCR)	80
9.3.2.	Colony-PCR.....	80
9.4.	Cloning Methods	81
9.4.1.	Restriction Digestion.....	81
9.4.2.	Ligation	81
9.4.3.	CloneJet™ PCR Cloning Kit.....	82
9.4.4.	Transformation of chemically competent cells.....	82
9.4.5.	Sequencing of genomic and plasmid-DNA.....	83
9.4.6.	PCR for hTERT and GFP	84
9.5.	Experiments with human cell lines.....	84
9.5.1.	Cultivation of human cell lines.....	84
9.5.2.	RNA isolation	84
9.5.3.	Reverse Transcription PCR (RT-PCR)	85
9.5.4.	Transient Transfection	85
10.	Results	86
10.1.	Construction of a telomerase-specific expression vector.....	86
10.1.1.	Deletion of the CMV-promoter	87
10.1.2.	Amplification of the hTERT-promoter sequence	87
10.1.3.	GFP amplification.....	88
10.2.	Cloning strategy	89
10.3.	hTERT-Expression	90
10.3.1.	Evaluation of telomerase-activity by qRT-PCR	90
10.3.2.	Testing the functionality of our new constructed telomerase-specific vector by transient transfection of different cell lines.....	91
11.	Discussion	95
11.1.	Evaluation of Telomerase-activity by qRT-PCR	96
12	Literature	98
13.	Supplementary	106

14.	Curriculum vitae	128
------------	-------------------------------	------------

1. Abstract

1.1. Abstract

With the growing number of available predictive biomarkers clinical management of cancer is becoming increasingly reliant on the accurate serial monitoring of tumor genotypes. In this thesis project it was tested whether changes in the tumor genome can be inferred from the peripheral blood of patients with cancer. To this end “liquid biopsies”, i.e. the analysis of circulating tumor DNA (ctDNA) and circulating tumor cells (CTCs) was of special interest. To elucidate the potential of ctDNA the plasma DNA of 32 patients with stage IV colorectal carcinoma (CRC) was analyzed. In a subset of the patients (34.4%) we found the presence of long (i.e. ~320 bps) DNA fragments, which were associated with increased CTC numbers and elevated concentration of mutated plasma DNA fragments. In these cases we were able to establish genome-wide tumor-specific copy number alterations directly from plasma DNA. When we analyzed plasma DNA from 35 patients with metastatic breast cancer we made similar observations, suggesting that our approach may be applicable to a variety of tumor entities. However, an unexpected finding was that not all patients with progressive metastatic disease appear to release tumor DNA into the circulation in measurable quantities. Furthermore, we sought to facilitate the detection of CTCs. These cells, shed from the primary tumor or metastatic lesions, co-occur with millions of cells from the hematopoietic system and are very rare. Therefore the detection and isolation of these cells pose a technological challenge and most of the currently used systems for CTC-detection lack specificity and sensitivity. We developed a telomerase-specific adenoviral-expression-vector, which is able to transfect blood cells. Due to the inserted expression-cassette - consisting of a human telomerase reverse transcriptase promoter (hTERT) linked to the green fluorescent protein (GFP) - GFP is expressed exclusively in telomerase-active cells. Tests with various tumor cell lines, such as the highly telomerase-active colorectal cancer cell line HCT 116, the weak telomerase-expressing breast cancer cell line MCF 7, and the telomerase-negative osteosarcoma cell line U2-OS, suggested that our constructed adenoviral-vector is able to mediate GFP-expression in telomerase-active cancer cells. Thus, it might be an applicable tool to identify CTCs with ease and to make them amenable for further analyses. In summary, we show in this thesis that we could establish the current copy number status of a tumor genome by non-invasive means from the peripheral blood. Furthermore, we developed a novel detection system, which may greatly facilitate detection of CTCs.

1.2. Zusammenfassung

Mit der zunehmenden Anzahl der mittlerweile verfügbaren prädiktiven Biomarker wird das klinische Vorgehen bei Patienten mit Krebserkrankungen immer mehr von der genauen Beobachtung der Tumor-Genotypen abhängig. In Zuge dieses Projekts versuchten wir festzustellen, ob aus dem peripheren Blut von Krebspatienten Tumor-spezifische Kopienzahl-Veränderungen nachzuweisen sind. Schließlich waren wir in der Lage, eine Größenverteilung der im Blut zirkulierenden Plasma-DNA zu identifizieren und den Anteil der mutierten Plasma-DNA mittels Deep-Sequencing und ultra-sensitiven Mutations-Detektions-Verfahren, z.B. BEAMing, zu bestimmen. Bei der Analyse der Plasma-DNA von 32 Patienten, die an einem Kolonkarzinom im Stadium IV leiden, zeigte sich, dass ein Teil der Patienten (34.4%) eine biphasische Verteilung der Plasma-DNA-Fragmente beinhaltete, welcher mit einer erhöhten Anzahl an zirkulierenden Tumor-Zellen (CTCs) im Blut und einem höheren Gehalt an im Blut zirkulierender Plasma-DNA korrelierte.

Bei diesen Patienten war die Ermittlung der genom-weiten Tumor-spezifischen Kopienzahl-Veränderung der Plasma-DNA direkt aus dem peripheren Blut der Patienten möglich. Der aktuelle Stand der Kopienzahl-Veränderungen konnte analysiert werden, und das, obwohl die Diagnose des primären Tumors in einigen Fällen schon Jahre zurück lag.

Erstaunlicherweise zeigten, trotz vorangeschrittener Erkrankung, nicht alle Patienten die Freisetzung von messbarer Tumor-DNA ins Blut.

Weiters ergab die Untersuchung der Plasma-DNA von 35 Patienten mit metastasierender Brustkrebs-Erkrankung sehr ähnliche Resultate, was die Spekulation zulässt, dass die von uns etablierte Methode ebenso für andere Tumor-Entitäten anwendbar ist.

In dieser Arbeit wird zum ersten Mal das Vorkommen einer Größenverteilung innerhalb der Plasma-DNA beschrieben, welche einen wichtigen Einfluss auf die Tumor-Diagnose und die Überwachung der Erkrankung haben könnte.

Das Ziel des zweiten Teils dieses Projektes war es, einen adenoviralen-Vektor zur Detektion von zirkulierenden Tumor-Zellen zu entwickeln. Diese zirkulierenden Tumor-Zellen (CTCs), werden vom primären Tumor und von den Metastasen abgegeben, vermischen sich mit den Komponenten des peripheren Blutes und sind aus diesem Grund nur in einer sehr geringen Anzahl vorhanden. Darum stellen sowohl die Detektion als auch die Isolierung der CTCs eine sehr große technische Herausforderung dar. Vor allem, weil es diesen Systemen meist an Spezifität und Sensitivität fehlt. Deshalb entschieden wir uns zur Entwicklung eines Vektors,

der alle Zellen des Blutes infiziert, jedoch – mit Hilfe der enthaltenen Expressions-Kassette, bestehend aus hTERT-Promoter und GFP – lediglich Tumor-Zellen zum Leuchten bringt.

Um die Transkriptions-Fähigkeit des entwickelten Adenovirus zu untersuchen transfizierten wir verschiedene Zelllinien mit dem Telomerase-Vektor: Die schwach Telomerase-aktiven Brustkrebszellen MCF 7, die stark Telomerase-aktiven Kolonkarzinom-Zellen HCT 116 und Osteosarkom-Zellen, die keine Telomerase exprimieren.

Die Ergebnisse der Transfektion zeigten, dass der hTERT-GFP-Vektor die GFP-Expression in Telomerase-aktiven Zellen vermitteln kann und daher eine nicht-invasive Möglichkeit zur Überwachung des Tumor-Wachstums und Voranschreiten der Erkrankung bieten könnte.

1.3. Abbreviations

°C	degree Celsius
ACF	aberrant crypt foci
ALT	alternative lengthening of telomers
ATP	adenosintriphosphat
Array-CGH	array-comparative genomic hybridization
BFB-cycle	breakage-fusion-breakage cycle
BUB	budding uninhibited by benimidazoles
bp	basepairs
CAM	cell adhesions molecule
CIMP	CpG island methylator phenotype
CIN	chromosomal instability
CRC	colorectal cancer
CTCs	circulating tumor cells
ctDNA	circulating tumor DNA
DAPI	4', 6-diaminodino-2-phenyliodole
DMEM	Dulbecco's Modified Eagle Medium
DNA	deoxyribonucleic acid
EGF	epidermal growth factor
EMT	epithelial-to-mesenchymal transition
FAP	familial adenomatous polyposis
Fig.	Figure
FITC	flouresceinisoithiocyanat
EpCAM	Epithelial cell adhesions molecule
FOBT	faecal occult blood testing
GAPS	GTPase activating proteins
GEFs	guanin exchange factors
GDP	guanosin diphosphat
GFP	green fluorescent protein
HCC	hereditary colon cancer
HEK	human embryonic kidney cells
HNPCC	hereditary non-polyposis colon cancer

ICC	Immunocytochemistry
IRES	ribosomal entry site
MAD	mitotic-arrest-deficient
Mb	mega base
MCS	multi cloning site
MIN	microsatellite instability
MMR	mismatch repair
MSI	microsatellite instability
NSAID	nonsteroidal anti-inflammatory drugs
PBS	phosphate buffered saline
PD	population doublings
RNA	ribonucleic acid
Supp.	supplementary
TP 53	tumor protein p53
WGA	whole genome amplification

Part I: Establishment of tumor-specific copy number alterations from plasma DNA of patients with cancer

Int J Cancer. 2013; 8(1); Doi:10.1002/ijc.28030

Establishment of tumor-specific copy number alterations from plasma DNA of patients with cancer

Ellen Heitzer^{1*}, Martina Auer^{1*}, Eva Maria Hoffmann^{1*}, Martin Pichler^{2*}, Christin Gasch³, Peter Ulz¹, Sigurd Lax⁴, Julie Waldispuehl-Geigl¹, Oliver Mauermann³, Sumitra Mohan¹, Gunda Pristauz⁵, Carolin Lackner⁶, Gerald Höfler⁶, Florian Eisner², Edgar Petru⁵, Heinz Sill⁷, Hellmut Samonigg², Klaus Pantel³, Sabine Riethdorf³, Thomas Bauernhofer^{2,8}, Jochen B. Geigl^{1#}, Michael R. Speicher^{1#}

1: Institute of Human Genetics, Medical University of Graz, Harrachgasse 21/8, A-8010 Graz, Austria

2: Division of Oncology, Medical University of Graz, Auenbruggerplatz 15, A-8036 Graz, Austria

3: Institute of Tumor Biology, University Medical Center Hamburg Eppendorf, Martinistr. 52, D-20246 Hamburg, Germany

4: Department of Pathology, General Hospital Graz West, Goestingerstrasse 22, A-8020 Graz, Austria

5: Department of Obstetrics and Gynecology, Medical University of Graz, Auenbruggerplatz 14, A-8036 Graz, Austria

6: Institute of Pathology, Medical University of Graz, Auenbruggerplatz 25, A-8036 Graz, Austria

7: Division of Hematology, Medical University of Graz, Auenbruggerplatz 38, A-8036 Graz, Austria

8: Present address: Thomas Bauernhofer, LKH Leoben, Department for Hematology and Oncology, Vordernberger Straße 42; A-8700 Leoben, Austria.

*: these authors contributed equally

Correspondence: jochen.geigl@medunigraz.at (J.B.G.), michael.speicher@medunigraz.at (M.R.S.); phone: +43 316 380 4110; fax: +43 316 380 9605

2. GENERAL INTRODUCTION

2.1. Colorectal Cancer

2.1.1. Epidemiology of colorectal cancer

Compared to all types of cancer, colon cancer is the third common (after breast cancer in women and prostate cancer in men) diagnosed type of cancer in Austria.

Over 1.2 million new cases and 608.700 deaths were reported in 2008 (Jemal *et al.*, 2011). Both among men and women incidence rates were nearly fourfold higher in more developed regions of the world compared to less developed (Kamangar *et al.*, 2006) and colorectal cancer incidence rates continued to increase in economically transitioning countries, whereas rates stabilized or even decreased in long-standing economically developed countries (Center *et al.*, 2009).

In Austria the rate of incidence of colorectal cancer among the population is 30 per 100.000, whereas the mortality is 13 per 100.000. Each year 4.500 new cases and 2.200 cancer deaths are reported. Compared to women, the incidence rate in men is increased by 1.7x and the rate of mortality by 1.8x. During the last century the rate of newly diagnosed colorectal carcinomas has decreased by 24% and the mortality by 28% (Statistik Austria, 2010).

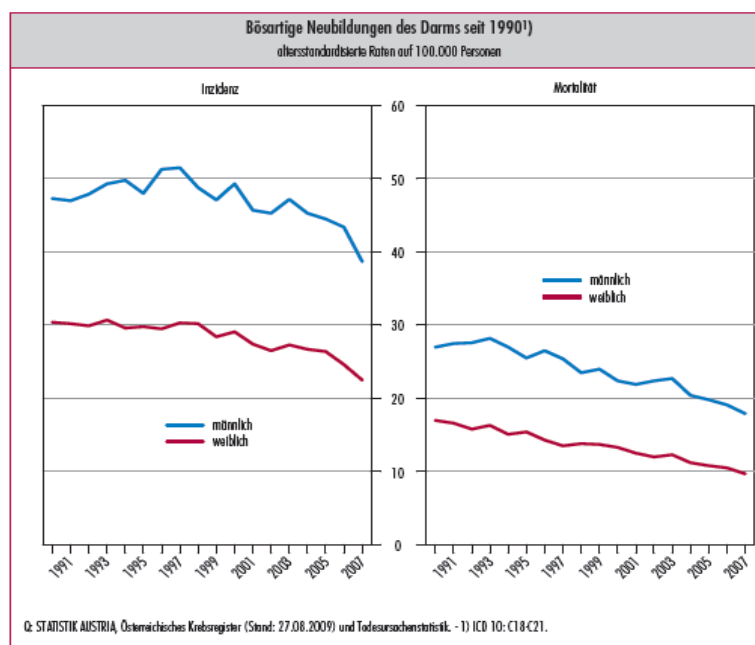


Figure 1: Colorectal Cancer, Incidence and Mortality in Austria

2.1.2. Stages of colorectal cancer

At present, the most commonly used staging system is the TNM (tumor/ [lymph] nodes/ metastases) system. Some TNM stages are summarized to broader stages and then often quoted as number I, II, III, or IV. The formerly frequently used Dukes classification has been largely replaced by the TNM system.

For example, stage 0 corresponds to the TNM stage TisN0M0, and is also termed carcinoma *in situ*, because abnormal cells are restricted to the mucosa. Stage I combines the TNM stages T1N0M0 and T2N0M0, summarizing tumors, which have invaded the submucosa and muscularis propria, respectively. The most advanced stage is IV where distant metastases are present (Ludwig & Weinstein, 2005).

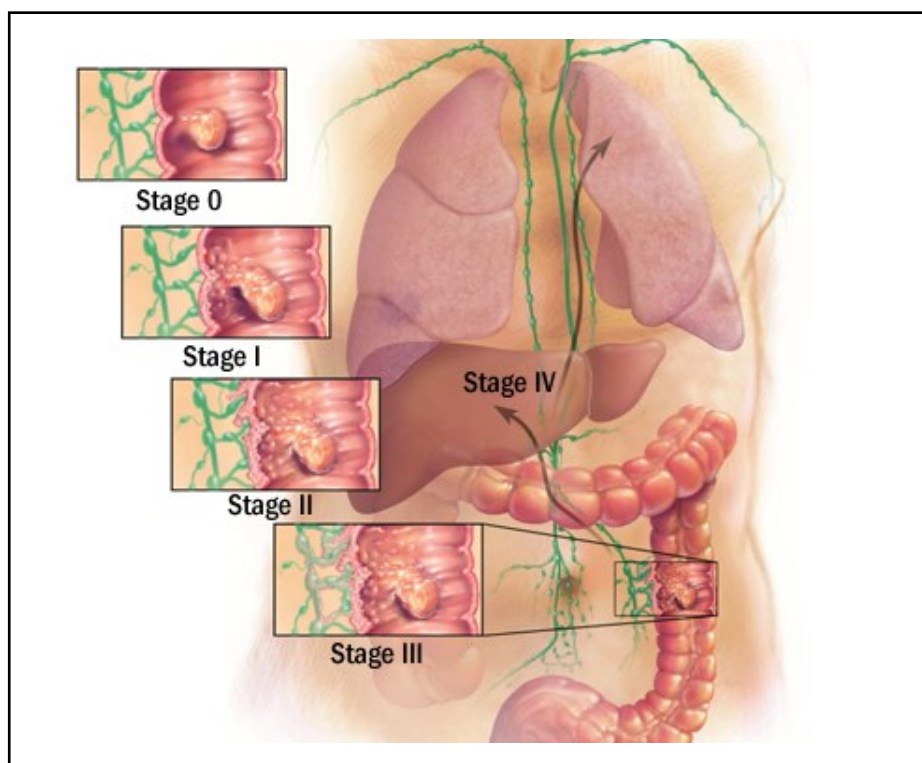


Figure 2: Stages of colorectal cancer: Colon cancer development from stage 0 until stage IV (Source: <http://www.mayoclinic.com/health/medical/IM01892>)

2.3. Cancer Metastasis

In order to form metastases a cancer cell has to spread from the primary tumor and undergo a series of steps (Chambers, 2002), termed **invasion-metastasis-cascade** (Valastyan & Weinberg, 2011; Chaffer & Weinberg, 2011).

In order to ensure the supply with metabolic substances and blood the tumor needs access to blood vessels in the course of tumor formation, a procedure called angiogenesis (Chambers, 2002; Geiger & Peeper, 2009).

In the first step of the metastatic process, cancer cells separate from the primary tumor, disseminate into the nearby tumor-associated stroma and into surrounding tissue. As long as the normal tissue remains well organized, the tumor cells are unable to overcome this barrier (Valastyan & Weinberg, 2011; Gupta & Massagué, 2006).

The process of intravasation is characterized by the entrance of cancer cells into lymphatic or new built blood vessels. For that reason, the passage through the endothelial wall of the microvessels is permitted by changes of the molecular composition (Valastyan & Weinberg, 2011; Chambers, 2002). These modified tumor cells can stimulate the formation of new instable and leaky blood vessels in order to simplify the passage for tumor cells into the lumina (Valastyan & Weinberg, 2011).

Cancer cells have two possible ways to disseminate from a primary tumor:

a.) hematopoietic route: tumor cells are disseminating from the primary tumor and build metastases in distant organs. This is the case in patients who develop metastases at distant sites, whereas lymph nodes are not affected. Hematopoietic dissemination is assumed to start in an early stage of cancer progression.

b.) lymphatic route: cancer cells disseminate from the primary tumor, pass to the lymph node, and form lymph node metastases. Once cancer cells are at distant sites they are dormant, even for several years.

After progress of the disease, tumor cells disseminate from the lymph node to distant sites and build secondary metastases (Pantel & Brankenhoff, 2004; Gupta & Massagué, 2006).

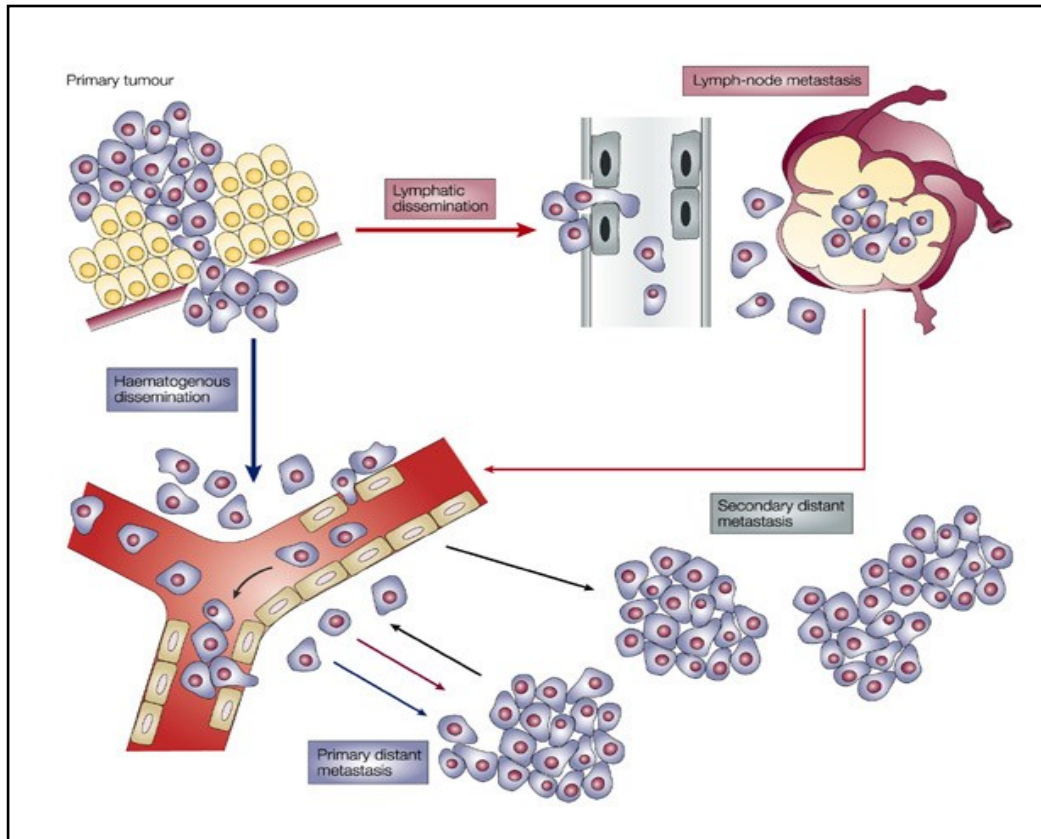


Figure 3: Cancer cell dissemination (Source: Pantel & Brankenhoff, 2004). Cancer cells have two possible ways to disseminate from the primary tumor: **a.) hematopoietic route** (blue), **b.) lymphatic route** (red). Black arrows indicate cancer cells disseminating from distant metastatic sites into the blood vessel.

In theory, circulating tumor cells exhibit the ability to disseminate to several distant sites, but it has been shown that different types of carcinomas metastasize to different distant sites. Two theories exist regarding **the arrest of CTCs at distant organs** (Valastyan & Weinberg, 2011).

- a.) The limited diameter of blood vessels displays a size limitation for circulating tumor cells. Therefore CTCs arrest within the capillary (Chambers, 2002).
- b.) It is hypothesized that CTCs are led to a specific distant site due to defined receptor-ligand-interactions, which would indicate an active homing process of these cells (Gupta & Massagué, 2006).

The process of **extravasation**, which is defined as the step where tumor cells leave the endothelial vasculature and move into target tissue, differs within several tumors (Gupta & Massagué, 2006). In order to pass the vascular barrier, tumor cells start to secrete factors destroying distant microenvironment and increasing vascular permeability. These secreted factors are for example COX-2, metallo-proteases (MMP-1, MMP-2) (Valastyan & Weinberg, 2011), and Src family kinases (Gupta & Massagué, 2006).

For a further development of **micrometastases** the extravasated CTC establish a “premetastatic niche” which is responsible that the new microenvironment shows a big difference from the primary tumor. The primary tumor releases signals for the mobilization of hematopoietic progenitors and sends them to a distant site in order to change the microenvironment and make it more hospitable.

The majority of circulating cells are unable to form clinically relevant metastases, because cells suffer slow attrition over a long period of time, or they are in a micro metastatic state and remain dormant (Chamber, 2002). In order to gain the ability for proliferation and to break through the dormancy, cell-nonautonomous processes are needed for the transformation of foreign microenvironments into convenient niches (Valastyan & Weinberg, 2011). The process of colonization is an inefficient event, because most of the cancer cells which translocate from the primary tumor to a distant site undergo apoptosis (Chaffer & Weinberg, 2011).

In 1889, Paget’s “seed and soil” theory was published in *The Lancet*. He hypothesized that although all tumor cells are able to disseminate, metastases can only develop at sites (“soils”) where disseminating tumor cells (“seeds”) proliferate and survive (Chambers, 2002; Valastyan & Weinberg, 2011).

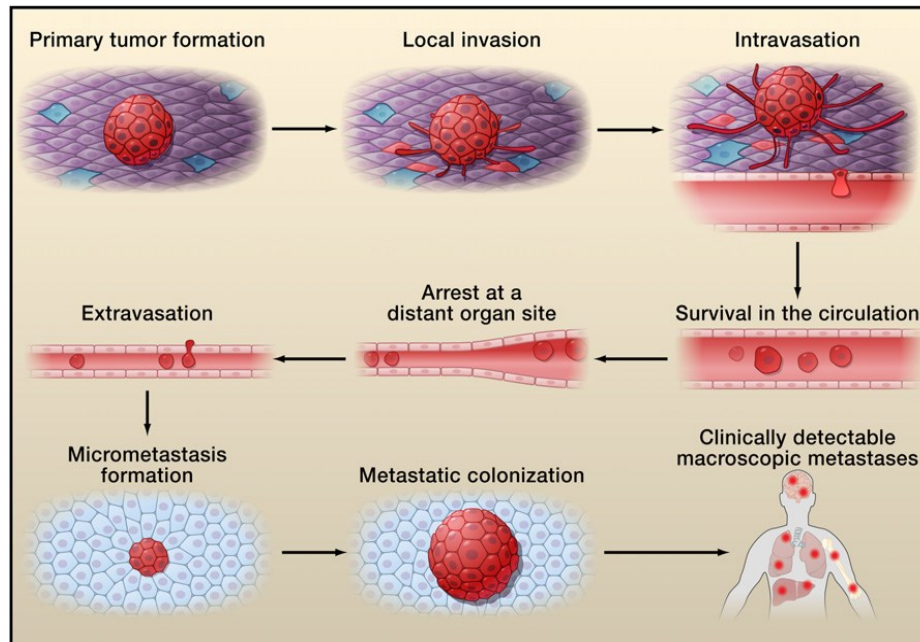


Figure 4: The metastatic progression (Source: Valastyan & Weinberg, 2011)

The aforementioned summary of the present knowledge about metastasis illustrates that many aspects of this important process are unknown. Therefore, the analyses of circulating tumor cells or of circulating tumor DNA from the peripheral blood of patients with cancer should

have great potential to learn more about metastases as outlined in the next sections of this thesis.

2.4. Circulating DNA

Five years before Watson and Crick defined the double helical structure of DNA, Mandel and Métais identified cell-free circulating DNA floating around in human blood plasma (Anker *et al.*, 1999). In 1977 Leon and Shapiro showed that cancer patients had higher concentrations of circulating cell-free DNA in plasma than healthy individuals, but due to technological limitations it was not possible to further investigate these nucleic acids (Leon & Shapiro, 1977). Circulating DNA in plasma was not only found in cancer patients, but was also present in the plasma of healthy individuals (Jung *et al.*, 2010) and in patients with other diseases, including pancreatitis, oesophagitis, hepatitis, stroke, and inflammatory bowel disease (Anker *et al.*, 1999). Another possible reason for an increased level of circulating DNA in plasma is excessive overtraining (Fatouros *et al.*, 2006).

During pregnancy a median of 10% of the DNA in the plasma is of fetal origin (Lun *et al.*, 2008). After liver and kidney transplantation donor derived DNA was detected in the plasma of recipients (Lo *et al.*, 1999).

2.4.1. Circulating tumor DNA (ctDNA) in cancer patients

Whereas the mean concentration of double stranded ctDNA in healthy individuals is about 10 - 20 ng/ml plasma (Jung *et al.*, 2010; Jahr *et al.*, 2001), the mean level of circulating plasma DNA detected in cancer patients was 180 ng/ml plasma. It has been shown that significantly higher DNA levels could be detected in plasma of patients with metastases (Jahr *et al.*, 2001; van der Vaart, 2008). Diehl *et al.* described that for a patient with a primary tumor with a weight of 100g (which corresponds to 3×10^3 tumor cells), approximately 3.3% of tumor DNA may enter the blood every day (Diehl *et al.*, 2005; Schwarzenbach *et al.*, 2011).

The reliable detection of ctDNA is a particular challenge due to the low amount of DNA in plasma, combined with the condition that ctDNA represents only a fraction of the total circulating DNA (Diehl *et al.*, 2005). Additionally, plasma DNA is highly fragmented (size distribution in the range of 180-10,000bp) (Jahr *et al.*, 2001). Concentrations of circulating DNA are influenced by cancer dependent factors, for example tumor size, stage, location, and other risk- and prognosis related parameters (Jung *et al.*, 2010).

Cancers arise through sequential alteration of genes that control cell growth and other cellular processes. It is known that in solid tumors approximately 80 genes harbor mutations which are present in tumor cells but not in normal cells (Wood, 2007). Furthermore, it has been demonstrated that ctDNA shows tumor-related alterations, e.g. point mutations, microsatellite alterations in renal carcinomas (Goessl *et al.*, 1998) and head neck cancer patients (Nawroz *et al.*, 1996) decreased strand stability (Anker *et al.*, 1999), and aberrant methylation of tumor suppressor genes (Fehm & Banys, 2011; Sanchez-Cespedes *et al.*, 2000). In 1998, Hibi *et al.* used mutations in *K-Ras* or *TP53* as a target for the detection of tumor-related DNA alterations in a serum of colorectal cancer patients (Hibi *et al.*, 1998;). In theory, these mutations could serve as potential biomarkers and could also be a useful tool for improving the detection and monitoring of the metastatic diseases (Diehl *et al.*, 2008). Especially for this purpose and for a better understanding of the disease non-invasive approaches would be highly attractive.

2.4.2. Circulating tumor DNA

Several studies to understand the origin of plasma DNA have been performed. According to the reported characteristics, a high percentage of cell-free circulating DNA may originate from degrading cells of the primary tumor (Jahr *et al.*, 2001).

Regarding the release of DNA into the blood circulation, three hypotheses are discussed:

1.) **Circulating tumor cells undergo lysis** and release DNA into circulation (Gormally *et al.*, 2007). This is controversially discussed, because a lot more circulating tumor cells should have been present in the blood (van der Vaart, 2008).

2.) Apoptosis and/or necrosis of tumor cells:

Nucleosomes bear a core protein complex consisting of a histone octamer, where DNA fragments with a length of 200bp are wound around. If a cell undergoes apoptosis the nucleosomes are packed into apoptotic particles and are going to be engulfed by phagocytes (macrophages and dendritic cells) (Choi *et al.*, 2005; Schwarzenbach *et al.*, 2011). Thereby the chromatin is cleaved into fragments with a length corresponding to multiple mono- and oligonucleosomes (180 – 200bp) by a set of activated caspases (van der Vaart, 2008; Nagata, 2005). This characteristic pattern is also seen after electrophoresis of cell-free circulating plasma DNA and indicates an apoptotic origin of circulating DNA in plasma (Jahr *et al.*,

2001; Atamaniuk *et al.*, 2006). In contrast to this, fragments larger than 10 000 bp are thought to originate from necrotic cells (Jahr *et al.*, 2001).

One fact against necrosis is that after the treatment of patients with radiotherapy a decrease in the cell-free circulating DNA level was observed. However, if necrosis was indeed the source of circulating DNA, an increase in cf-DNA concentration should have been the consequence (Leon & Shapiro, 1977).

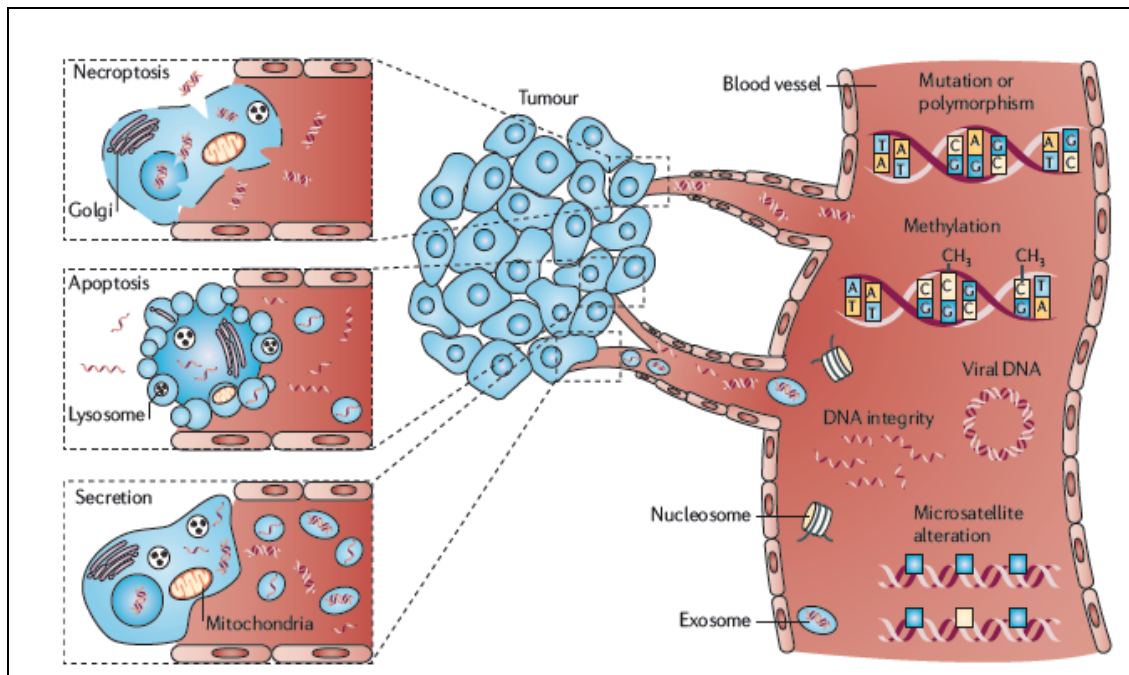


Figure 5: Cell-free circulating DNA in the blood. (Source: Schwarzenbach *et al.*, 2011). Circulating DNA in plasma can be detected due to several mutations, different methylation patterns, or microsatellite alterations. Tumor-related circulating DNA enters the blood stream after apoptosis/necrosis, but it can also be actively released into the circulation.

3.) Release of DNA into the bloodstream:

In healthy individuals it has been shown that circulating DNA is derived from hematopoietic origin. Lui *et al.* performed quantitative real-time PCR to estimate the amount of Y-chromosomal DNA in the plasma of patients after sex-mismatched bone-marrow transplantation, and they found that circulating cell-free DNA was predominantly of donor/hematopoietic origin (Lui *et al.*, 2002).

2.4.3. The role of ctDNA as biomarker in carcinogenesis

During the last decades, various trials of plasma DNA analysis in persons with cancer have been attempted. Most reports have focused on region-specific tests, e.g. detection of mutations which had previously been identified in the corresponding primary tumor from the same patient.

The reason for these attempts is that DNA which contains somatic mutations is highly tumor-specific and thus provides, in theory, a much more specific indicator of neoplasia than any other biomarker (Diehl *et al.* 2008a, 2008b).

Muller and colleagues used microsatellite analysis in order to detect allelic shifts in the plasma DNA (Muller *et al.*, 2008) and investigated specific patterns of methylation within plasma DNA fragments (Gormally *et al.*, 2007). Although these studies have succeeded in defining cell-free circulating tumor DNA, the approaches were not sensitive enough to serve as a diagnostic tool.

In 2005, Diehl and colleagues established a method to quantify the amount of mutant and normal DNA in given plasma samples. This technique, called BEAMing (beads, emulsion, amplification, magnetics), enables investigators to detect a fraction of mutant DNA even at a very low level (Diehl *et al.*, 2005). Diehl *et al.* (2005) used an extension of BEAMing, which included in a first step determination of the total number of mutated circulating fragments by RT-PCR. Next, amplified oligonucleotides are transported into magnetic beads. Subsequently, the mutational status is evaluated by single nucleotide extension, and finally flow-cytometry is used to measure the FITC, Cy5 and phycoerythrin signals of the beads simultaneously (Diehl *et al.*, 2005). Especially the monitoring of mutations previously identified in the primary tumor has been suggested as a tool to improve the detection and monitoring of the metastatic diseases (Diehl *et al.*, 2008b).

Currently, novel sequencing technologies to identify alterations in the primary tumor genome are being employed. The presence of such alterations was analyzed in the plasma DNA and correlated with the tumor burden (Leary *et al.*, 2010; McBride *et al.*, 2010). However, sensitive and specific detection of a mutated base in a vast excess of normal DNA requires specialized techniques currently beyond the scope of many diagnostic laboratories (Diehl *et al.*, 2005; 2008a; 2008b; Yung *et al.*, 2009). Moreover, all of these technologies are extremely laborious, time-consuming, and expensive and have therefore not been introduced for real clinical applications yet. Once an eligible approach has been developed, cell-free circulating

tumor DNA could serve as “liquid biopsy” and could be further useful for multiple diagnostic applications (Schwarzenbach *et al.*, 2011)

2.5. Circulating tumor cells (CTCs)

Circulating tumor cells (CTCs) detach from the primary tumor, migrate in the blood circulation to distant organs, and are able to form distant metastasis. The occurrence of metastases is the reason for most cancer-related deaths (Yu *et al.*, 2011; Smirnov *et al.*, 2005). CTCs are thought to act as a potential biomarker and might provide significant prognostic and predictive information for treatment and guidance of monitoring of the disease (Dotan *et al.*, 2009). Therefore, these circulating cells attracted high interest in the field of cancer research (O’Flaherty *et al.*, 2012). Detection and isolation of CTCs from peripheral blood of cancer patients are technically challenging procedures. Currently only limited information regarding the numbers of CTCs in the blood stream at different stages and types of cancer is available. Also their heterogeneity and their significance in the natural history of cancer disease are not well understood (Yu *et al.*, 2011). However, CTCs could give important information about the molecular and functional conditions of the primary tumor, and provide further insight into cancer biology. The approach of CTCs as “liquid biopsy” can serve as a basis for the identification of novel therapeutic markers for treatment and prevention of blood-borne cancer progression (Maheswaran & Haber, 2010, O’Flaherty *et al.*, 2012). Right now there is no gold standard in the current approaches for the optimal CTC isolation strategy (Yu *et al.*, 2011). Several isolation methods are based on **physical properties** of the CTCs to distinguish tumor cells from normal cells in the peripheral blood of cancer patients. These characteristics are: larger size from epithelial cell, different density, migration properties, and specific protein expression of different cell types (Yu *et al.*, 2011; Dotan *et al.*, 2009). The isolation of circulating tumor cells by cell size is performed by micro porous polycarbonate filters (O’Flaherty *et al.*, 2012). In order to separate CTCs and mononuclear cells from other blood cells, density gradient centrifugation (Ficoll-Hypaque, Oncoquick) is employed (Sun *et al.*, 2011). The most commonly used strategy is based on using antibodies against cell surface-antigens. Since circulating tumor cells originate from tumors with epithelial origin, **epithelial surface markers** could be used for their detection (Yu *et al.*, 2011; Sleijfer *et al.*, 2007). However, dissemination of tumor cells into the blood stream is also associated with epithelial-to-mesenchymal-transition (EMT) so that CTCs may lose their epithelial character (Sun *et al.*, 2011; Sleijfer *et al.*, 2007). The process of EMT is described by loss of cell-to-cell adhesions,

accompanied by a more relaxed cytoskeleton and an increased expression of mesenchymal markers such as vimentin and *N*-cadherin (Sun *et al.*, 2011; Sleijfer *et al.*, 2007). This means that detection methods targeting tumor cells using epithelial surface markers may miss the most invasive tumor cells (Sun *et al.*, 2011). Nevertheless, the most commonly used approach for CTC detection is based on the expression of EpCAM (epithelial cell adhesion molecule), which is not present in normal blood cells (Maheswaran & Haber, 2010). To date, the FDA (Food and Drug Administration)-approved method, the CellSearchTM System (Veridex), is widely used (Yu *et al.*, 2011). The system enriches EpCAM expressing cells with antibody-coated magnetic beads. For specific identification of epithelial cells cytokeratin 8, 18, and 19 are recognized by two phycoerythrin-conjugated anti-cytokeratin antibodies. White blood cells are stained with an anti-CD45 antibody conjugated with allophycocyanin in order to differentiate them from epithelial cells. Further on, cells are counterstained with DAPI (4', 6-diamidino-2-phenylindole) for identification of intact cells (Riethdorf *et al.*, 2007; Cristofanilli *et al.*, 2004; O'Flaherty *et al.*, 2012). The CellSearchTM System identifies CTCs as nucleated cells, expressing epithelial markers (cytokeratins), but lacking CD45 antigen on their surface (Dotan *et al.*, 2009). Other approaches for CTC detection based on epithelial surface markers are the "CTC-chip" (Nagrath *et al.*, 2007), the "herringbone- (HB) chip" (Stott *et al.*, 2010), flow-cytometry based methods (Cruz *et al.*, 2005), and the MagSweeper® (a magnetic stir bar) (Talasaz *et al.*, 2009). Some additional innovative CTC detection strategies are microscope scanning approaches like the fibre-optic array scanning technology, laser-scanning cytometry, multiphoton intravital flow cytometry, photoacoustic flowmetry, and the so called CAM (Cell Adhesion Molecule) coated slides or vials (Yu *et al.*, 2011).

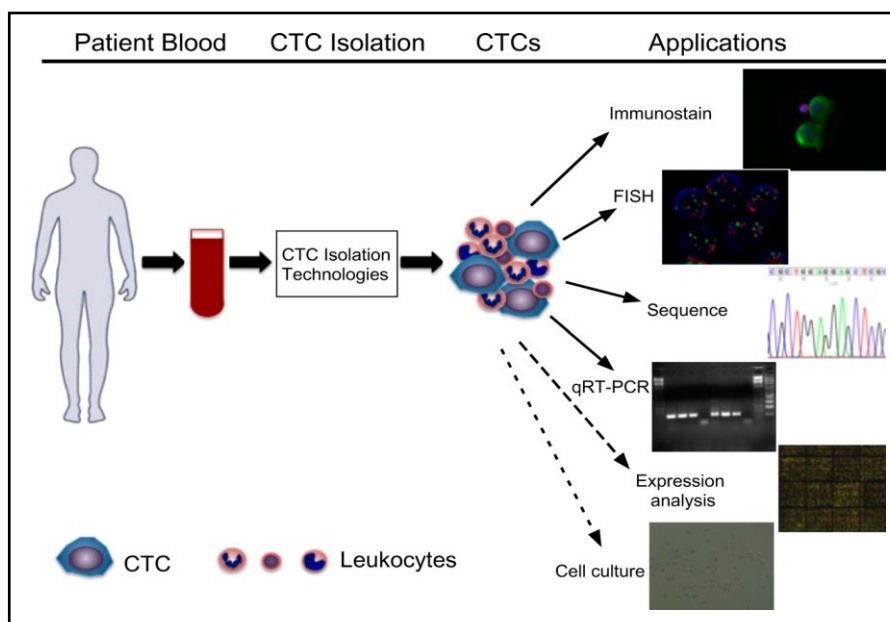


Figure 6: Illustration of current and potential applications of CTC technologies (peripheral blood of a cancer patient is collected and different CTC isolation technologies are performed) (Source: Yu *et al.*, 2011).

Isolated cancer cells can be further analyzed by techniques based on immunocytochemistry (ICC) or by reverse-transcriptase polymerase chain reaction (RT-PCR) for the characterization of nucleic acids (Dotan *et al.*, 2009). These approaches are based on the detection of markers which are thought to be more cancer specific. However, due to the heterogeneity of tumors, the use of these markers is problematic and no “universal” tumor marker is available at the moment (Sleijfer *et al.*, 2007). RT-PCR based-methods are applicable for CTC detection, but enumeration of cells and comparison between patients is not feasible (Dotan *et al.*, 2009; Sleijfer *et al.*, 2007).

Several studies have concluded that the presence of detectable CTCs in the blood serves as an independent prognostic factor in patients with breast, prostate, and colon cancer (Yu *et al.*, 2011).

2.6. Tumor monitoring and liquid biopsies

Besides studying the biology of metastasis, the analyses of ctDNA and CTCs also offers the opportunity to monitor changes in the tumor genome. This is outlined in the following (taken from our publication by Heitzer *et al.*, 2013):

Recent advances in the understanding of the molecular mechanisms of cancer have highlighted the need for personalized medicine approaches not only in terms of prognosis, but also for diagnostic strategies. Extensive work has resulted in the identification of biomarkers that have been implemented in cancer clinical practice at several levels: prognostics, predictive, and pharmacokinetic biomarkers. For example, in colorectal cancer (CRC) the *KRAS* mutations in exon 2 (codons 12 and 13) represent a paradigm that has been established as a negative predictive marker for treatment with epidermal growth factor receptor (EGFR) inhibitors such as cetuximab and panitumumab (Walther *et al.*, 2009). The discovery of further biomarkers will be accelerated by the recent advances in cancer genomics.

Access to accurate and sensitive methods for the detection of such predictive biomarkers is of utmost importance to the clinical oncologist. Numerous studies have tried to identify such biomarkers in the form of circulating DNA, because tumors shed DNA into the circulation, which can be detected by appropriate means. However, early reports proposing that the presence or absence of circulating DNA or its concentration is of diagnostic value (Leon *et al.*, 1977, Stroun *et al.*, 1989) have been called into question (Pinzani *et al.*, 2010, Schwarzenbach *et al.* 2011, Sidransky, 2002). Analyses of plasma/serum DNA for loss of heterozygosity (LOH) (Muller *et al.*, 2008, Schwarzenbach *et al.*, 2009) or for tumor-related methylation patterns (Begum *et al.*, 2011, Gormally *et al.*, 2007) often lack specificity. Furthermore, it has been proposed that the detection of mutations in peripheral blood, which have previously had identified in the corresponding primary tumor from the same patient, may provide a specific biomarker of disease burden. Therefore multiple studies focused on the detection of such specific and predetermined mutations (Diehl *et al.*, 2005; Diehl *et al.*, 2008a; Diehl *et al.*, 2008b; Leary *et al.*, 2010, McBride *et al.*, 2010; Nawroz *et al.*, 1996, Yung *et al.*, 2009). For example, the emergence of *KRAS* mutant clones as evidence for acquired resistance to targeted EGFR blockade in patients with colorectal cancer has been inferred from the analysis of plasma DNA (Misale *et al.*, 2012; Diaz *et al.*, 2012). A broader approach is the use of targeted amplicon sequencing for the simultaneous analysis of multiple cancer driver genes (Forsheew *et al.*, 2012). However, the sensitive and specific detection of a mutated base in a vast excess of normal DNA requires specialized techniques that are currently beyond the scope of many diagnostic laboratories (Diehl *et al.*, 2005; Diehl *et al.*, 2008a,b; Yung *et al.*, 2009).

In this study we investigated whether a genome-wide view of tumor-specific copy number changes instead of a targeted approach can be established from peripheral blood, i.e. plasma

DNA from patients with cancer. To this end we simultaneously quantified the normal and mutant DNA molecules in a given sample and established genome-wide copy number changes from plasma and compiled detailed data on copy number changes in relation to the respective primary tumors and metastases. We validated these observations with blood samples from 35 patients with breast cancer. Our results suggest that complex tumor genomes can be reconstructed from the peripheral blood of patients with cancer. Our approach may be of interest for two scenarios: first, as a research tool to investigate a tumor genome at late stage disease, which is not commonly studied. Second, it may pave the way for new disease monitoring strategies.

3. AIM OF THE THESIS

3.1. Specific aims:

The **main hypothesis** of this part of the thesis was that it may be possible to reconstruct tumor genomes out of blood from patients with colorectal cancer.

The advantage of such an approach would be the monitoring of complex tumor genomes from peripheral blood in a non-invasive way. For this purpose cell-free circulating tumor-DNA, as well as material from primary tumor/metastases (if available) and CTCs, were analyzed. After isolation of cell-free circulating DNA and evaluation of the DNA-concentrations qualitative and quantitative analysis were performed. Plasma DNA-samples were also used for a high-resolution array-CGH analysis in order to evaluate tumor specific chromosomal aberrations. Additionally, primary tumor/metastatic material was analyzed and compared with occurring aberrations within plasma DNA.

3.1.1. The following research questions were defined:

Is it possible to monitor whole tumor genotypes from peripheral blood of colorectal cancer patients with high-resolution molecular techniques?

What amount (percentage) of cell-free circulating plasma DNA is needed to detect genomic aberrations of tumors in peripheral blood of CRC patients and are these aberrations corresponding to the genomic profile of CTCs analyzed with high-resolution array-CGH (comparative genomic hybridization)?

Does the size distribution of cell-free circulating DNA correspond to the detection rate of tumor specific aberrations?

Is there a correlation between the number of CTCs in the peripheral blood of colorectal cancer patients and the size distribution of the cell-free circulating-DNA?

Is cell-free circulating DNA suitable to serve as diagnostic parameter for colorectal cancer patients?

4. Material and Methods

4.1. Material

4.1.1. Chemicals and Media

Acetonitrile, Emplura	Sigma Aldrich, Chemie GmbH, Germany
Agar Agar	Carl Roth GmbH & Co., Karlsruhe, Germany
Albumin from bovine serum	Sigma Aldrich, Chemie GmbH, Germany
Ammoniumchlorid	Carl Roth GmbH & Co., Karlsruhe, Germany
Ampicillin, Natrium Salz (Amp)	Serva Feinbiochemica, Heidelberg, Germany
ATP	Roche, Germany Holding GmbH
Aqua bidest.	Promega Germany GmbH, Mannheim
Bacto Trypton	Difco Laboratories, Detroit, MI, USA
Bacto Yeast extract	Difco Laboratories, Detroit, MI, USA
BigDYE Terminator v3.1	Applied Biosystems, Darmstadt, Germany
10 x aCGH Blocking Agent	Agilent Technologies, Austria
Chloroform, p.a.	Carl Roth GmbH & Co., Karlsruhe, Germany
Cy-dyed dCTP	GE Healthcare, Vienna, Austria
DMEM –	
Dulbecco's Modified Eagle Medium	Invitrogen, Vienna, Austria
dNTP Set (dATP, dCTP, dGTP, dTTP)	Roche, Germany Holding GmbH
DNA loading Dye	MBI Fermentas, St. Leon-Roth, Germany
Ethanol absolute, p.a.	Merck KGaA Darmstadt, Germany
Gasket-slides	Agilent Technologies, Austria
Gel Red Nucleic Acid Gel Stain	VWR, Vienna, Austria
HotStartTay Polymerase Mastermix	Qiagen GmbH, Hilden, Germany
Human Cot-1 DNA	Roche, Germany Holding GmbH
Human Cot-1 DNA	Invitrogen, Vienna, Austria
Human genomic DNA	Promega GmbH, Mannheim, Germany
2 x Hi-RPM Hybridization Buffer	Agilent Technologies, Austria
Isoamylalcohol	Carl Roth GmbH & Co., Karlsruhe, Germany
LE Agarose	Biozym, Vienna, Austria
LiChrosolv water	Merck KGaA Darmstadt, Germany
8 x 60K human Microarray	Agilent Technologies, Austria

3M NaAC, pH 5,2	Sigma Agilent Technologies, Austria
Neutral buffered formalin	Sigma Agilent Technologies, Austria
Nuclease-free water	Promega GmbH, Mannheim, Germany
Oligo aCGH Wash Buffer 1 and 2	Agilent Technologies, Austria
OnePhorAll Buffer	Amersham Biosciences Europe GmbH, Germany
Opti-MEM® Reduced Serum Medium	Invitrogen, Vienna, Austria
Peptone from Casein	Merck KGaA Darmstadt, Germany
Primer	Microsynth, Switzerland
Primer	Metabion GmbH, Martinsried, Germany
Proteinase K	Qiagen GmbH, Hilden, Germany
5 x Sequencing Buffer	Applied Biosystems, Darmstadt, Germany
Sodium Chloride	Merck KGaA Darmstadt, Germany
Yeast extract	Carl Roth GmbH & Co., Karlsruhe, Germany

4.1.2. Kits and Standards

DNA standards

100 bp DNA ladder	Invitrogen, Vienna, Austria
1 kb ladder	Qiagen GmbH, Hilden, Germany
Gelpilot wide range ladder	Qiagen GmbH, Hilden, Germany
Gene Ruler™ 1 kb DNA ladder Mix	MBI Fermentas, St. Leon-Roth, Germany

DNA Isolation Kits

GenElute™ Plasmid Miniprep Kit	Sigma Agilent Technologies, Austria
QiaAMP DNA Mini Kit	Qiagen GmbH, Hilden, Germany
QiaAMP Circulating Nucleic Acids Kit	Qiagen GmbH, Hilden, Germany

DNA Purification Kits

GenElute™ PCR CleanUp Kit	Sigma Agilent Technologies, Austria
GenElute™ Gel Extraction Kit	Sigma Agilent Technologies, Austria
Microcon YM-30	Millipore GmbH, Schwalbach, Germany
QiaAMP DNA Micro Kit	Qiagen GmbH, Hilden, Germany
Sephadex G-50	GE Healthcare, Vienna, Austria

DNA Labelling Kit

BioPrime Array-CGH Genomic
Labeling Module

Invitrogen, Vienna, Austria

DNA Amplification Kits

Illustra GenomiPhi V2 DNA
Amplification Kit

GE Healthcare, Vienna, Austria

Repli-g Mini Kit

Qiagen GmbH, Hilden, Germany

WGA-Kit (WGA 2)

Sigma Agilent Technologies, Austria

WGA-Kit (WGA 4)

Sigma Agilent Technologies, Austria

DNA Polymerase Kits and Enzymes

AdenoXpression System 1

Clontech, Takara, Japan

Alkaline Phosphatase

MBI Fermentas, St. Leon-Roth, Germany

CloneJET™ PCR Cloning Kit

MBI Fermentas, St. Leon-Roth, Germany

Fast Start HighFidelity PCR System Roche, Germany Holding GmbH

Fugene HD, Transfection reagent

Promega GmbH, Mannheim, Germany

Phusion High-Fidelity DNA

Polymerase

Finnzymes, Finland

Polymerase LA Taq™

Takara, Japan

Restriction Enzymes

New England Biolabs, Frankfurt, Germany

SYBR-Green Mastermix

Applied Biosystems, Darmstadt, Germany

Taq DNA Polymerase

New England Biolabs, Frankfurt, Germany

T4-Ligase

New England Biolabs, Frankfurt, Germany

Quant-iT™ Pico Green dsDNA

Assay Kit

Invitrogen, Austria

Quantitative Telomerase Detection Kit

Biocat, Heidelberg, Germany

4.1.3. Bacterial strains

***E.coli* DH5alpha**, (Invitrogen, Vienna, Austria): F⁻ Φ 80*lacZ* Δ M15 Δ (*lacZYA-argF*) U169 *recA1 endA1 hsdR17* (rK⁻, mK⁺) *phoA supE44* λ ⁻ *thi-1 gyrA96 relA1*

***E.coli* Top10**, (Invitrogen, Vienna, Austria): F⁻ *mcrA* Δ (*mrr-hsdRMS-mcrBC*) Φ 80*lacZ* Δ M15 Δ *lacX74 recA1 araD139* Δ (*ara leu*) 7697 *galU galK rpsL* (Str^R) *endA1 nupG*

4.1.4. Plasmids

Name	Antibiotic Resistance	Company
pShuttle2	Kan ^r	Clontech
pShuttle2/hTERT	Kan ^r	
pShuttle2/hTERT/GFP	Kan ^r	
pAcGFPN1	Kan ^r	Clontech
pJet 1.2	Amp ^r	MBI Fermentas

4.1.5. Cell lines

Name	Description	Media
BJ (PD40)	human fibroblasts	DMEM (Gibco BRL)
CaCO2	human colorectal adenocarcinoma	DMEM (Gibco BRL)
HCT116	human colorectal carcinoma	DMEM (Gibco BRL)
Hep3b	human hepatocellular carcinoma	RPMI (Gibco BRL)
HepG2	human hepatocellular carcinoma	RPMI (Gibco BRL)
HT29	human colorectal adenocarcinoma	McCoy's 5a (Gibco BRL)
MCF7	human breast adenocarcinoma cell line	DMEM (Gibco BRL)
SKBR3	human breast adenocarcinoma cell line	McCoy's 5a (Gibco BRL)
SW480	human colorectal adenocarcinoma	Leibovitz's L-15 (Gibco BRL)

Media was supplemented with 10% Gold-FBS (PAA) and 1% Penicillin/Streptomycin (PAA).

4.1.6. Oligonucleotides

Sequencing Primers:

Vector/hTERT_F	5' - ttggaggtcgctgagtagtgc-3'
hTERT_3_R	5' - cagcgctgcctgaaactc-3'
hTERT/GFP_R	5' -tggtgcagatgaacttcagg-3'
hTERT/GFP_F	5' - ccccttcacctccagctc-3'
GFP_1_F	5' - ccggtcatggtgagcaag-3'
GFP_1_R	5' - gtatcgccctcgaacttcac-3'
GFP_2_F	5' - atcacatgaagcagcacgac-3'
GFP_2_R	5' - tagtggttatcgggcagcag-3'
GFP_3_F	5' - acttcaagatccgccacaac-3'
GFP_3_R	5' - tcggtatggtgtaaacgaat-3'

Oligonucleotides used for restriction site attachment:

hTERT_ <i>MfeI</i> Fwd	5' - aaacaattgagtggattcgcgggcacaga-3', (Song, 2004)
hTERT_ <i>XbaI</i> Rev	5' - aaatctagaagggtctccacgtgcgcag-3', (Song, 2004)
GFP_ <i>NotI</i> Fwd	5' - taagcggccgcaccggtcatggtgagcaag-3'
GFP_ <i>AflIII</i> Rev	5' - caacttaagcttgtagctcatccatg-3'

Oligonucleotides used for qRT-PCR:

hTERT_qPCR_F	5' - acagcctgtttctggatttg-3'
hTERT_qPCR_R	5' -catgcgtgaaacctgtacg-3'
HPRT_qPCR_F	5' -acactggcaaaacaatgcag-3'
HPRT_qPCR_R	5' -gtggggctccttttcaccag-3'

Primers SRY-determination:

SRY_F:	5' -tggcgattaagtcaaattcgc-3'
SYR_R:	5' -ccccctagtaccctgacaatgtatt-3'
Beta-globin_F:	5' -gtgcacctgactcctgaggaga-3'
Beta-globin_R:	5' -ccttgataccaacctgccag-3'
GAPDH_F:	5' -ccccacacacatgcacttacc-3'
GAPDH_R:	5' -cctagtcccagggtttgatt-3'

4.1.7. Laboratory equipment

Bioanalyzer 2100	Agilent Technologies, Austria
Centrifuge FRESCO 21	ThermoScientific
Concentrator Plus	Eppendorf, Germany
Cuvettes	Promega GmbH, Mannheim, Germany
Electrophoresis Chambers	
7500 Fast Real-Time PCR System	Applied Biosystems, Darmstadt,
Germany	
Flourescence Imaging VILBER	PEQLAB, Erlangen, Germany
Hybridisation oven	Shel Lab, USA
Incubator	Heraeus Instruments, Germany
Microarray Scanner with Surescan High-	
Resolution Technology	Agilent Technologies, Austria
Rotina 48 R	Hettich, Germany
Orbital incubator SI500	Stuart, UK
Power Pac	BioRad, Germany
QuantiFlour™	Promega, Madison, WI, USA
Thermocycler, DNA Engine	BioRad, Germany
Thermomixer	Eppendorf, Germany
Ultraviolet Sterilizing PCR Workstation	PEQLAB, Erlangen, Germany
Vacuum pump	Millipore GmbH, Schwalbach,
	Germany
Vortexer	IKA, Staufen, Germany
Waterbath	Lauda, Julabo, GFL Germany

Pipettes

Gilson P10, P20, P100, P1000	Abimed Analysen-technik GmbH,
	Germany
Pipettes 5 ml, 10 ml, 25 ml, 50 ml	Greiner Bio-One International AG,
	Austri

Tubes

Sterile tubes 15ml, 50ml

Greiner Bio-One International AG,
Austria

Safe-Lock Tubes 1,5ml, 2ml

Eppendorf, Germany

PCR Softtubes 200 µl

Biozym, Vienna, Austria

PCR-Tubes Maximum Recovery

Axygen, California, USA

Glas flasks

100 ml, 500 ml, 1000 ml

Schütt Labortechnik GmbH, Germany

Erlenmeyerkolben

50 ml, 100 ml

Schott Duran GmbH, Germany

4.1.8. Buffers

LB-media

Trypton 10 g/l

Yeast extract 5 g/l

Sodium chloride 5 g/l

For preparation of agar plates:

Agar Agar 20 g/l

10x TE

12,11 g Tris

3,72 g Na₂EDTA pH 8,0

Stop-Solution

50 mg Bromphenolblue, 17,24 ml glycerin 87% 242,20 mg EDTA to 50 ml bidestilled water.

4.1.9. Software, Database

Vector NTI

Invitrogen, Vienna, Austria

Agilent Scan Control Program

Agilent Technologies, Austria

Feature Extraction

Agilent Technologies, Austria

4.2. Methods

4.2.1. Sample collection

4.2.1.1. Blood samples after sex-mismatch bone marrow transplantation

All patients (n=20) received a sex-mismatch bone marrow transplantation after acute myelogenous leukemia (AML). Patients were seen at the Division of Haematology, Department of Internal Medicine from the Medical University of Graz. The study was approved by the local ethics committee and written informed consent was obtained from all patients.

Twelve out of these twenty patients were male and had received bone marrow transplantation from a female donor, whereas eight patients were female and had received bone marrow transplantation from a male donor.

Control blood samples of six healthy individuals were collected.

Peripheral blood of patients was collected in EDTA VAcutainer® tubes (BD Biosciences, Heidelberg, Germany) and plasma was prepared according to Dhallan *et al.* (2004).

4.2.1.2. CRC-Patients

(These data are published in Heitzer *et al.*, 2013). All patients with colon cancer (n=32) had advanced stage (UICC Stage IV), progressive disease at the time of blood collection. Clinical evidence of disease progression was based on (a) newly diagnosed metastatic disease, (b) disease progression after initial treatment response during treatment break, or (c) primary tumor progression as initial response to anti-tumor treatment. Twenty-three patients were male, 9 female, and the mean age at diagnoses was 68 years (range: 45-81 years) (Supp. Table 1).

The intervals between diagnosis of the primary tumor and our blood collection, between diagnosis of the primary tumor and metastasis, and between blood collection and last therapy are listed in Supp. Table 1. The interval between diagnosis of the primary tumor and blood collection to obtain both plasma DNA and circulating tumor cells (CTCs) varied extensively within a range from 1 month (patients #2, #9, #18, #23) to more than 12 years (151 months; patient #28) (mean: 32 months; median: 24 months) (Supp. Table 1). The mean interval between blood collection and last therapy for all patients together was 104 days (median: 47 days; range: 0-688 days).

In addition to the clinical data in the main text we provide here some data regarding patients #9, #33, and #38:

Patient #9: At the time of diagnosis metastases were present in liver, bone, and abdomen (peritoneal carcinomatosis). We obtained biopsies from the primary tumor and the metastasis and performed our analyses from the peripheral blood one month after diagnosis.

Patient #33: At the time of diagnosis metastases were present in liver and abdomen (peritoneal carcinomatosis and abdominal wall metastasis). The primary tumor was completely removed and a liver biopsy was obtained four months later; our analyses were done 9 months after initial diagnosis.

Patient #38 had the primary tumor completely resected at diagnosis. Forty-four months later metastases were noted in liver, bone, and lung; a biopsy was obtained from the liver. One month later we performed our analyses.

The most frequent sites of metastases were liver (diagnosed in 25 patients), followed by lung ($n=11$), peritoneal carcinomatosis ($n=8$) and bone ($n=6$). Supp. Table 3 summarizes how often we identified the biphasic plasma DNA size distribution in these patients and how many CTCs were found.

In patients #6 and #9 the primary tumor was only biopsied, in patients #26, #33, and #38 the primary tumor was resected completely. In three of these patients (i.e. #6, #9, #33) metastases were already present at the time of diagnosis of the primary tumor, in the remaining patients (i.e. #26 and #38) the interval between diagnosis of the primary tumor and metastasis was 31 and 44 months, respectively. At the time of our blood collection all of these patients had metastases at several different sites (Supp. Table 1).

The study was approved by the local ethics committee and written informed consent was obtained from all patients. The colon cancer patients were seen at the Division of Clinical Oncology, Department of Internal Medicine, at the Medical University of Graz. The blood from patients with repeated blood drawings on 5 consecutive days was obtained from the LKH Leoben. The breast cancer patients were treated at the Department of Obstetrics and Gynecology, Medical University of Graz.

4.2.2. Plasma preparation

(published in Heitzer *et al.*, 2013). Whole blood (9 ml) was collected in routine EDTA Vacutainer tubes (BD Biosciences, Heidelberg, Germany). To stabilize cell membranes

and to impede cell lysis 0.225 ml of a 10% neutral buffered solution containing formaldehyde (4% weight per volume) (Sigma-Aldrich, Vienna, Austria) was added immediately after blood withdrawal. Blood samples were gently inverted, stored at room temperature, and further processed within two hours. Plasma was prepared according to Dhallen *et al.*, 2004.

In brief, tubes were centrifuged at 200 g for 10 min. with the brake and acceleration powers set to zero, followed by a subsequent centrifugation step at 1600 g for 10 min. The supernatant was collected, transferred to a new 15 ml tube, and spun at 1600 g for 10 min. The plasma was carefully aliquoted into new 2 ml Eppendorf tubes and stored at -80°C . DNA was isolated from plasma samples using the QIAamp DNA Blood Mini Kit (Qiagen, Hilden, Germany) or the Qiagen Circulating Nucleic Acids Kit (CNA) (Qiagen, Hilden, Germany) according to the manufacturer's instructions. DNA was eluted in 30 μl of distilled water for Mini Kit extractions and in 100 μl for CNA Kit, respectively.

4.2.3. Isolation of cell-free circulating DNA from plasma

4.2.3.1. QIAamp Mini Kit

Cell-free circulating DNA was isolated from plasma using the QIAamp DNA Mini Kit (#51304, Qiagen, Hilden, Heidelberg) according to the manufacturer's instructions (DNA Purification from Blood or Body Fluids). In brief, 200 μl plasma was mixed with Proteinase K (1/10 of the plasma volume), 200 μl AL buffer, incubated for 10 min at 56°C . DNA was precipitated with 200 μl of Ethanol (100 %), before loaded on the column containing a silica membrane. During a centrifugation step ($6000 \times g$ 1 min) negatively charged DNA molecule could bind on the positively charged silica membrane whereas unbind molecules were washed from the column in two centrifugation steps ($20000 \times g$). Plasma DNA was eluted from the silica membrane using 30 μl distillate water (preheated to 55°C), incubated for 1 min and centrifuged at $6000 \times g$. Plasma DNA was stored at -20°C .

4.2.3.2. QIAamp Circulating Nucleic Acids Kit

The plasma DNA was isolated by the QIAamp Circulating Nucleic Acids Kit (Qiagen, Hilden, Heidelberg) according to the manufacturer's instructions (Purification of circulating nucleic acids from 1 ml, 2 ml or 3 ml serum or plasma).

Cell-free circulating DNA was isolated from 2 ml plasma. In order to release plasma DNA from proteins, lipids and vesicles and for inactivation of DNases and RNases the sample was lysed under highly denaturing conditions. For that incubation with 200 µl Proteinase K and Buffer ACL at (60°C, 30 min) was needed. Adding 3.6 ml ACB buffer to the lysate allowed optimal binding conditions of DNA fragments to the silica membrane.

Lysates were transferred to QIAamp Mini columns onto silica membranes and drawn through by vacuum pressure. Nucleic acids bound on the membrane and unbind molecules were washed away with three additional steps.

To evaporate the Ethanol from the silica membrane the column was dried for 10 min at 56°C and in a single step pure DNA was eluted in 100 µl distilled water, equilibrated at room temperature.

Samples were directly used for downstream applications or stored at -20°C.

4.2.4. Plasma-DNA analysis

4.2.4.1. Quantitative analysis of plasma DNA

4.2.4.1.1. *Quant-iT™ Pico Green dsDNA Assay*

(published in Heitzer *et al.*, 2013)

For quantification of cell-free DNA in plasma we used Quant-iT™ PicoGreen® dsDNA Kit (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions. We used bacteriophage lambda DNA for a standard curve (ranging from 500 pg/µl to 16 pg/µl). 50 µl of the standard dilutions were transferred into disposable cuvettes and mixed with 50 µl of Quant-iT™ PicoGreen® reagent working solution. Experimental DNA samples were diluted in 1xTE (1:25) to a final volume of 50 µl and also mixed with 50 µl of Quant-iT™ PicoGreen® reagent working solution in disposable cuvettes. Fluorescence of the samples was then measured in duplicates using a QuantiFluor™ Fluorometer (Promega, Madison, WI, USA). The fluorometer was calibrated using the blank and the highest concentration of the standard dilutions. DNA concentration of the samples was determined from the standard curve generated in Excel. From these DNA amount estimates we used 800 pg for a qualitative bioanalyzer analysis and, if available, 50 ng for the whole-genome amplification.

4.2.4.1.2. Nanodrop

DNA concentration was determined by NanoDrop® ND- 1000 spectrophotometer (PeqLab, Germany).

The system was blanked by the use of water or 1/10 elute buffer. 1.5 µl of the sample were pipetted onto the end of a fiber optic cable. Due to a contact with a second fiber optic cable the liquid sample built a bridge between fibre optic ends. The sample is assessed at both a 1 mm and 0.2 mm path, this ensures an extensive dynamic range for the measurement of ds DNA. (Quantitation of Nucleic Acids Using A Microvolume Spectrophotometer Without The Use Of Cuvettes Or Capillares, Current Protocols in Molecular Biology). A xenon flash lamp provides the light source and the DNA concentration is determined by the absorption from 260 nm.

Additionally the absorption of the sample at a wavelength of 280 nm was measured to ensure that the sample was not contaminated by proteins. The ratio of 260/280 nm is in the range of 1.8–2.0 for high quality DNA.

4.2.4.2. Qualitative analysis of plasma DNA

4.2.4.2.1. Bioanalyzer 2100 Agilent (Agilent protocol)

In order to determine the length of DNA fragments in the plasma we used the Agilent 2100 Bioanalyzer. The minimal amount of plasma DNA from healthy subjects reflecting a peak in the Bioanalyzer profile was 800 pg. Hence, we used the same amount of plasma DNA to analyse the fragment size of native plasma DNA from cancer patients.

Agilent 2100 Bioanalyzer chips are produced for single use from Caliper Technologies Corp. (Mountain View, CA) and compounded of soda lime glass. The sample is analyzed via electrophoreses by loading the DNA on a compact chip-system (Agilent High Sensitivity Kit). Smaller fragments run faster than larger ones and they were compared with two internal standards to estimate DNA concentration and size.

For sample analysis, the DNA ladder and the gel have to be put into the wells of the chip. The wells are connected to each other by glass micro channel system. The microchip is inserted in the supplied priming station. 9 µl of a special gel-dye-mix are filled into the well marked with a black “G” via pressurization. After 1 min the priming station can be opened and the other “G” marked wells were filled with 9 µl of the gel. Into the wells for

sample analysis 5 µl of the marker and 1 µl of the probe were added. In order to monitor the size of the fragments 1 µl of DNA ladder was added into the corresponding slot.

The chip was vortexed for 1 min at 2400 rpm with a special mixer, directly inserted into the Bioanalyzer 2100 and analyzed with the program “ds DNA, High Sensitivity Kit”.

Due to fluorescence dye which intercalates into DNA the fluorescent signal is detected and translated into gel like images and electropherograms. The Agilent 2100 Bioanalyzer Expert software automatically determine size (bp), size distribution in CV (%), concentration (pg/µl), % of total and molarity (pmol/µl) of each defined smear region.

The Agilent 2100 Bioanalyzer (Agilent Technologies) allows sophisticated smear analysis for DNA sizing and quantification. In each sample we focused on the range from 80 to 600 bp, as we invariably observed merely a flat line without measurable fluorescence intensities below 80 bp or above 600 bp, indicating that no measurable DNA is present within this range. We applied sliders within the electropherogram to determine the percentage of the total DNA within the respective smear region. The sliders were positioned at the beginning and the end of the apoptosis peak and, if applicable, also at the second peak.

4.2.4.2.2. Real-time quantitative PCR

In our AML study we were interested to estimate the amount of Y chromosomal DNA from a male donor in the blood of a female recipient.

The protocol for SRY detection was described elsewhere (Lui *et al.*, 2002).

Therefore a quantitative Real-time PCR analysis was performed using an Applied Biosystems 7500 Fast Real Time PCR System.

Plasma DNA was amplified using Power SYBR®Green PCR Master Mix (Applied Biosystems), an intercalating dye that intercalates into double stranded DNA. As long as the fluorescent dye is free, it is undetectable. Once the dye is intercalating to dsDNA it starts to emit fluorescence light (Giulietti *et al.*, 2001). The polymerase chain reaction (PCR) product can be detected directly, because the fluorescence generated is proportional to the amount of product that is present in the sample. Quantitation based upon measurement of the increase in fluorescence signal of the sample with each cycle and comparison with a standard curve. (Horsman *et al.*, 2006)

Experimental setup

For detection of Y-chromosomal DNA following procedure was used.

Samples for qRT-PCR were prepared on ice and under the hood to ensure sterile conditions. 5 µl of isolated plasma DNA were amplified in duplicates using 50 µl from the reaction mix as follows:

Tab.1: Mastermix for qRT-PCR 1

Mastermix	1 x
Plasma DNA	5 µl
Power SYBR® Green	25 µl
Primer I (10µM)	2 µl
Primer II (10µM)	2 µl
water	16 µl

Polymerase chain reaction was performed according to the SYBR-Green protocol.

Tab. 2: qRT-PCR temperature program

I	Denaturation	95°C	10 min
II	Amplification (40x)		
	Denaturation	95°C	15 sec
	Annealing	60°C	1 min
III	Melting curve	95°C	15 sec
		60°C	1 min
		95°C	30 sec
IV	Cooling	60°C	15 sec

5 µl of water instead of plasma DNA were used as a negative control in these experiments. All patients' samples were amplified according to the previously described protocol.

Melting curve analysis

The major disadvantage of the SYBR green is the binding to any kind of ds DNA such as primers or new synthesized DNA by DNA polymerase. The melting temperature describes the temperature, whereas half of the DNA template is in a double stranded state. This is influenced by CG content and length of the sequence. The principle of melting curve

analysis is based on the release of fluorescent dye while DNA double strands are breaking which results in a decrease of fluorescence.

Primers for SRY and GAPDH were designed according to a publication from Lo *et al.*, 1998. Primers were diluted in distilled water to a concentration 10 μ M. Genomic DNA from three female and three male donors were used as a template and diluted to 100 ng, 10 ng, 1 ng and 0.1 ng. qRT-PCR was performed with PCR programme as described above.

Primer efficiency

Prior sample analysis, the primer efficiency has to be determined. In theory, the optimal primer efficiency is 2, meaning that during each PCR cycle a doubling of DNA takes place. In the practical implementation this value is usually not achieved. The value can either exceed (due to primer dimers or measurement inaccuracies) or be below (development of PCR by products).

In order to determine the primer efficiency Real-Time PCR was performed according to the protocol and serial diluted genomic DNA was used as a template.

Ct values obtained from each DNA dilution were used for calculation of primer efficiency (E) (Pfaffl *et al.*, 2001).

In this case the CT values were plotted against the logarithm of DNA concentration, resulting in a straight line with the slope (f) which was needed for further calculation of the efficiency using the formula:

$$E = 10(-1/f)$$

Calculation of Y-chromosomal DNA

In Order to determine the DNA concentration and for further calculation of the amount of Y-chromosomal DNA, the specific Ct-value of the each sample is needed.

The Ct (cycle threshold) is defined as the number of cycles required for the fluorescent signal to cross the exceeded background level. Ct levels are inversely proportional to the amount of target nucleic acid in the sample, meaning the lower the Ct level the higher is the target nucleic acid in the sample.

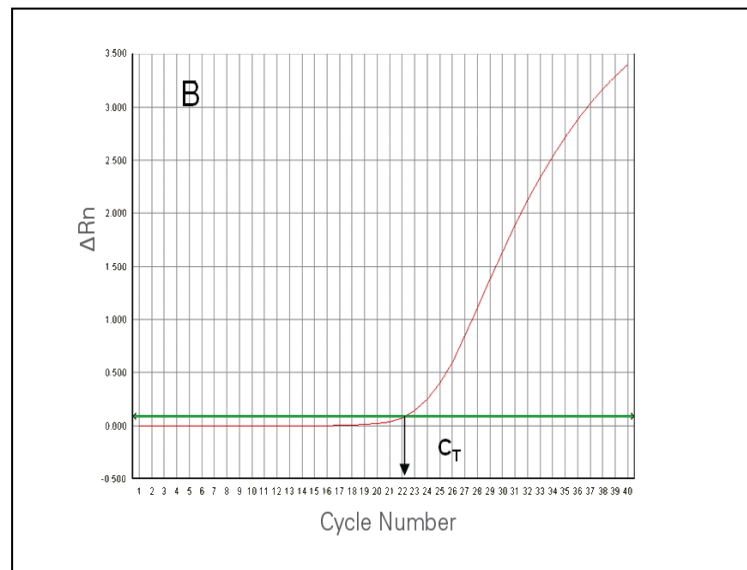


Figure 7: Ct- threshold cycles

(Source:http://www3.appliedbiosystems.com/cms/groups/mcb_marketing/documents/generaldocuments/cms_053906.pdf)

The Ct is the intersection between an amplification curve and a threshold line. For further quantification, a calibration curve was prepared. Genomic male DNA was diluted to concentrations of 66 ng, 6 ng, 660 pg, 66 pg, 6.6 pg and duplicates were run in parallel. Identical thermal profiles were used for both SRY and Beta Globin assay (Lo *et al.*, 1998). The amount of 6.6 pg of DNA (1 genome equivalent) represents the DNA content of one single cell and was used for expression of results as copy numbers.

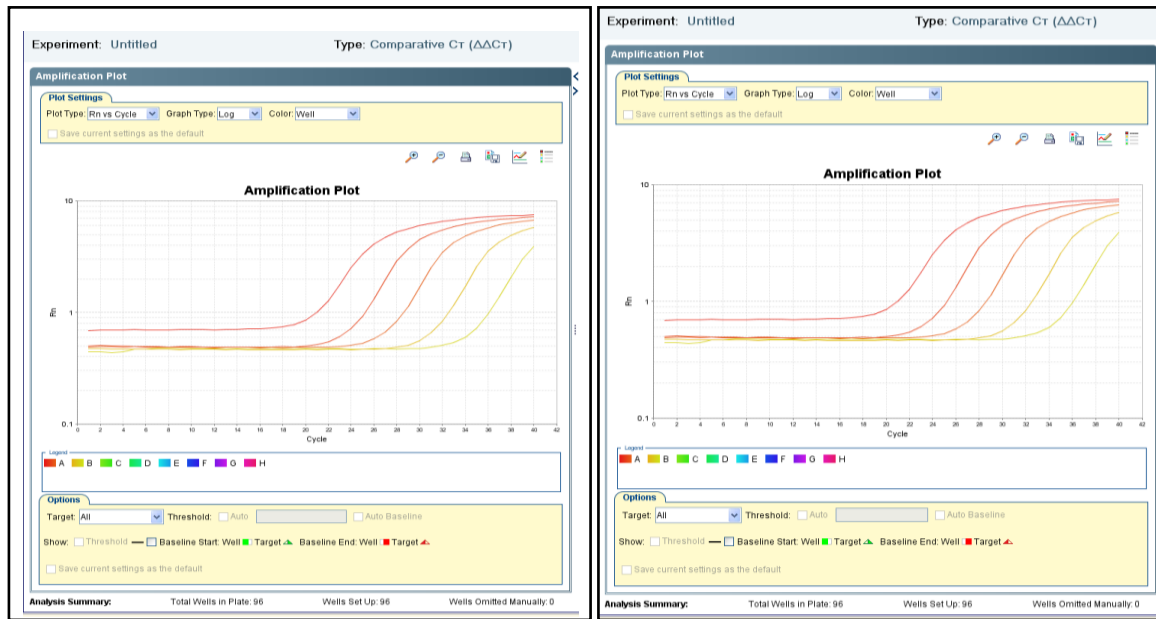


Figure 8: Amplification Plot for 10000 GE, 1000 GE, 100 GE, 10 GE, 1GE. a) SRY gene. b.) Beta-Globin

The mean quantity of each duplicate was used for further calculation.

The concentration, expressed in copies/ml, was calculated by use of following equation: (Lo *et al.*, 1998)

$$C = Q \times \frac{V_{\text{DNA}}}{V_{\text{PCR}}} \times \frac{1}{V_{\text{ext}}}$$

Because the SRY gene is found only in male nucleated cells, whereas the β -Globin gene is present in males and females, we calculated the percentage of male DNA in particular plasma, denoted as Y%, using the following equation: (Lui *et al.*, 2002)

$$Y\% = \frac{\text{SRY}}{\beta\text{-globin}} \times 100\%$$

4.2.5. DNA Isolation Methods

4.2.5.1. DNA Isolation of tumor DNA from FFPE sections

DNA isolation was performed by Martina Auer. (Data published in Heitzer *et al.*, 2013)

Formalin-fixed paraffin embedded (FFPE) tissue samples from primary tumors and, if available, from metastases were cut and mounted on a microscope slide. Hematoxylin and eosin stained slides were reviewed by experienced pathologists (C.L. and S.L.) and areas with a high tumor cell infiltration were macrodissected from parallel sections. In addition, DNA was isolated using the QIAamp DNA FFPE tissue kit (Qiagen, Hilden, Germany), following the manufacturer's instructions. Quantification and quality of the extracted DNA were determined using the Nano-Drop Spectrometer ND-1000 (Peqlab Biotechnologie, Erlangen, Germany).

4.2.6. DNA Amplification Methods

4.2.6.1. Amplification of cell-free tumor DNA

(Method description published in Heitzer *et al.*, 2013).

We adjusted the PCR procedure to minimize any further degradation of DNA and to allow an unbiased amplification of all fragments irrespective of their size. For our purposes, i.e. amplification of plasma-DNA with fragments of different lengths, it turned out that the GenomePlex Complete Whole Genome Amplification Kit (WGA2, Sigma-Aldrich, Vienna, Austria) is - with some modifications - best suited.

DNA concentrations of all samples were adjusted to the same content. We used 50 ng of plasma DNA as a template for further amplification. A total volume of 9 µl of DNA was used to prepare libraries according to the protocol. For that 1 µl of 10 x Fragmentation buffer, 1 µl Library Preparation Solution and 2 µl Library Preparation Buffer were added to DNA. In contrast to the original Sigma protocol we do not use the fragmentation step for plasma DNA amplification.

The mixture was vortexed thoroughly, consolidate by centrifugation and place in thermal cycler at 95°C for 2 min. Sample was cooled on ice, consolidate by centrifugation. 1 µl of Library Preparation Enzyme was added and the mixture was incubated at 16°C for 20 min, 24°C for 20 min, 37°C for 20 min, 75°C for 5 min and 4°C for library preparation.

Amplification was performed adding 7.5 µl of 10 x Amplification Master Mix, 47.5 µl of nuclease-free water and 5 µl WGA DNA Polymerase.

DNA Samples were amplified using an initial denaturation of 95°C for 3 min followed by 14 cycles, each consisting of a denaturation step at 94°C for 15 s and an annealing step at 65°C for 5 min.

4.2.7. Detection of CTCs and whole genome amplification

4.2.7.1. CellSearch System

The detection of circulating tumor cells and subsequent single-cell amplifications has already been published (Heitzer *et al.* 2013). Experiments were performed from our collaborators from the Institute of Tumor Biology, Medical University Eppendorf, Hamburg.

Blood samples (7.5 ml each) were collected into CellSave tubes (Veridex, Raritan, NJ, USA). The Epithelial Cell Kit (Veridex) was applied for CTC enrichment and enumeration with the CellSearch system as described previously (Riethdorf *et al.* 2007). In brief, in a first step CTCs were captured by anti-epithelial cell adhesion molecule (EpCAM)-antibody-bearing ferrofluid. Subsequent identification of CTCs was based on cytokeratin (CK)-positivity and negativity for the leukocyte common antigen CD45. In addition, 4',6-diamidino-2-phenylindole (DAPI) staining was done to evaluate the integrity of the nucleus.

For further analysis of CTCs, the EpCAM-enriched cell fraction in a volume of 300 µl was transferred from the CellSearch cartridge onto slides. Nucleated CK-positive, CD45-negative CTCs were isolated with the help of a micromanipulation device comprising the microinjector CellTram vario and the micromanipulator TransferMan NK2 (Eppendorf AG Hamburg, Germany) connected to an Axiovert 200 inverted microscope (Carl Zeiss AG, Jena, Germany). Single CTCs were placed into a 2.5 µl water drop in a 200 µl PCR reaction tube and stored overnight at -20°C prior to whole genome amplification. To propagate DNA from single CTCs, either the GenomePlex Single Cell Whole Genome Amplification Kit (Sigma-Aldrich, St. Louis, MO, USA) as described previously (Fiegler *et al.* 2006, Geigl *et al.* 2007) or the GenomiPhi DNA amplification kit (GE Healthcare, Chalfont St. Giles, UK) was applied.

For WGA using the GenomiPhi DNA amplification kit, 9 µl sample buffer and 1 µl protease (Qiagen, Hilden, Germany, dissolved in distilled water, final concentration 10.7 µAU/µl) were added for cell lysis (50°C for 15 min, 70°C for 15 min). The DNA was denatured (95°C for 2 min) and 10 µl reaction mix (9 µl reaction buffer + 1 µl enzyme

mix) was added. The amplification reaction was performed at 30°C for 2.5 h, followed by a final inactivation step at 65°C for 10 min.

WGA products were purified using the GenElute PCR Clean-up Kit (Sigma-Aldrich, St. Louis, MO, USA) for DNA obtained with the GenomePlex Single Cell WGA Kit (Sigma-Aldrich) or NucleoSeq Columns (Macherey-Nagel, Düren, Germany) for GenomiPhi amplified DNA. The final DNA concentration was determined by a NanoDrop ND-100 Spectrometer (Peqlab Biotechnologie, Erlangen, Germany).

When present in the primary tumor, we also tested for *KRAS* mutations in the CTC amplification products by Sanger sequencing.

4.2.8. DNA Purification Methods

4.2.8.1. GeneElute™PCR Clean Up Kit

WGA2 Amplification products were purified using the GenElute™ PCR Clean-up Kit (#NA1020; Sigma-Aldrich, Germany).

This method is developed for rapid purification of single stranded or double stranded PCR amplified products with a size from 100 bp to 10 kb from other components such as primers, dNTPs, DNA polymerases or salt.

500 µl of Column Preparation solution were used to prepare the silica membrane within the spin column for DNA binding. Amplified DNA was mixed with 5 volumes (350 µl) of Binding Solution, transferred to the silica membrane and columns were centrifuged at 12 000- 16 000 x g (room temperature). The bound DNA is washed with 500 µl wash solution centrifuging the columns for 1 min at 12 000- 16 000 x g and the columns were dried with an additional step of centrifuging (2 min 12 000- 16 000 x g). Amplified DNA was eluted with 50 µl of preheated nuclease free water and concentrations evaluated by Nanodrop.

4.2.8.2. QIAamp DNA Micro Kit

GenomiPhi amplified DNA was purified using the QIAamp DNA Micro Kit (Qiagen, Heidelberg, Germany). This kit is designed for cleaning RNA and high molecular weight DNA, such as genomic DNA. For purification DNA is bound on silica based membrane, washed and eluted with water.

The volume of the amplification product was exceeded to 100 µl with water. 10 µl buffer AW1 and 250 µl AW2 were mixed with the DNA. The sample was loaded on a column and centrifuged for 1 min at 6000 x g. The flow-through was discarded and the column was transferred to a new collection tube.

The column was washed with 500 µl Buffer AW 2 at 6000 x g for 1 min.

Then it has been moved to a new collection tube was dried by centrifugation for 3 min at 20 000 x g.

Finally the column was transferred to a new 1,5 ml tube, DNA was eluted from the membrane with 30 µl nuclease free water by centrifugation at 20 000 x g for 1 min.

4.2.9. DNA Labeling Methods

4.2.9.1. Bioprime® Exo-Knewlow labeling

DNA labeling was performed with the BioPrime® Array-CGH Genomic Labeling System. For that, random primer and a modified Knewlow fragment of DNA polymerase I for labeling of genomic or amplified DNA with fluorescently labelled nucleotides were used. Cy3 was used to label reference-DNA and Cy5 to label test-DNA.

Sample DNA concentration was determined by NanoDrop. For plasma DNA analysis 8 x 60k arrays were used.

250 ng DNA in 10.5 µl water and 10.5 µl 2.5 x Random Primer were denaturated for 5 min at 95 °C and immediately cooled on ice for 5 min. To this mixture 2.5 µl 10 x dCTPs, 1.5 µl Cy3 or Cy5 and 0.5 µl Exo-Knewlow were added. Incubation was performed for 2 hours at 37°C. The enzyme was inactivated at 65°C for 10 min.

Labeled DNA was either purified immediately by Amicon 30kDA columns or stored at -20°C.

4.2.9.2. Amicon 30kDA filters (Millipore)

The labeled samples were filled-up to a volume of 480 µl with 1 x TE (pH adjusted to 8.0) and loaded on the Amicon filters. Columns were centrifuged for 10 min at 14 000 × g in a microcentrifuge at room temperature. The flow- through was discarded. Again 480 µl of 1 x TE were loaded on the filter and columns were centrifuged for 10 min at 14 000 x g. The flow- through was discarded, the filter inverted and transferred to a new tube. In the following centrifugation step, the labeled DNA was collected in a new tube. After that the

sample volume is approximately 21 µl. For hybridization on an 8 x 60k array platform the sample volume should not exceed 10 µl. For that reason the sample were concentrated to 10 µl with a “concentrator plus” Speed-Vac centrifuge (Eppendorf, Germany).

1.5 µl of the sample were used to determine DNA concentration via NanoDrop® for further calculation of the yield, degree of labeling and specific activity of the sample.

4.2.9.3. Determination of the yield, the degree of labeling and specific activity

NanoDrop® used 1.5 µl of the labeled sample to determine the absorbance at A260 nm (DNA), A550 nm (cyanine 3), and A650 nm (cyanine 5).

The degree of labeling was calculated using following equation:

$$\text{Specific Activity}^* = \frac{\text{pmol per } \mu\text{L dye}}{\mu\text{g per } \mu\text{L genomic DNA}}$$

*pmol dyes per µg genomic DNA

Calculation of incorporation rate of the dye:

Yield [µg] = DNA concentration [ng/µl]*sample volume

Specific activity [pmol/µg] = (Dye [pmol/µg]*sample volume)/ Yield

Values expected after a successful labeling: (Agilent protocol)

Input gDNA (µg)	Yield (µg)	Specific Activity of Cyanine-3 Labeled Sample (pmol/µg)	Specific Activity of Cyanine-5 Labeled Sample (pmol/µg)
0.2 to 0.5*	2.5 to 3	15 to 25	15 to 20
0.5	5 to 7	25 to 40	20 to 35
3.0	7 to 10	35 to 55	25 to 40
* Half labeling reaction for 8x microarrays (half the amount of random primers, dye, enzyme and dNTPs)			

4.2.10. Comparative Genomic Hybridization, CGH

The molecular cytogenetic method of Comparative Genomic Hybridization is used to detect gains or losses (copy number variations), but also genomic imbalances in the human genome.

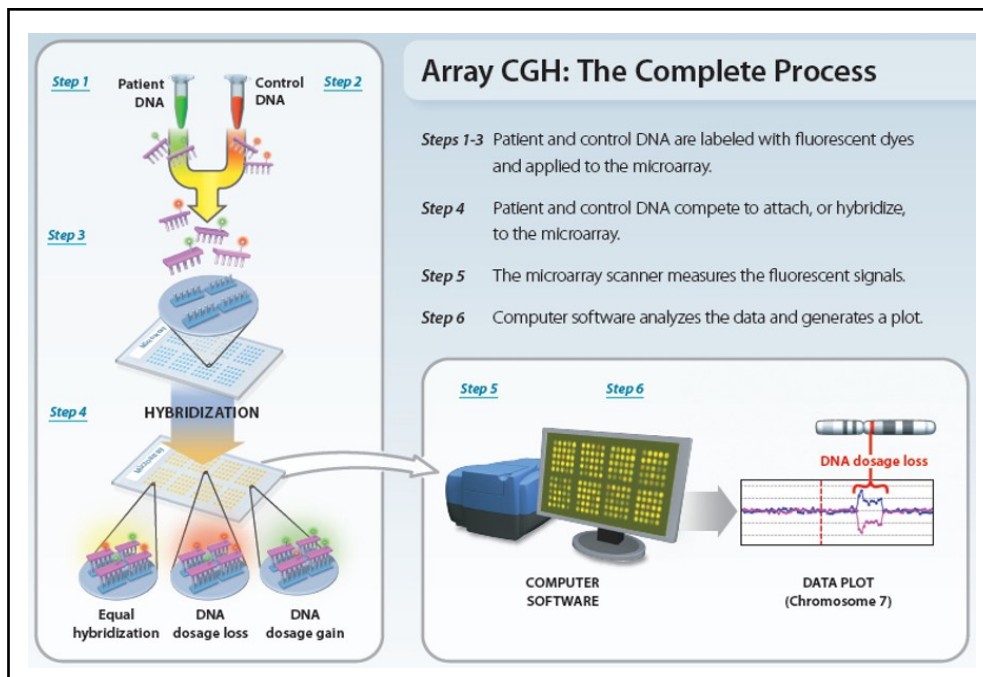


Figure 9: Array-CGH-method

(Source: http://www.nature.com/scitable/resource?action=showFullImageForTopic&imgSrc=/scitable/content/41020/arrayCGH_Theisen_FULL.jpg)

4.2.10.1. Array-CGH

(Description published in Heitzer *et al.*, 2013).

Array-CGH was carried out using a genome-wide oligonucleotide microarray platform (Human CGH 60k microarray kit, Agilent Technologies) protocol version 6.0. There, the whole genome is represented by oligonucleotides in the length of 60 bp (“Agilent 60-mer SurePrint technology”). The resolution of this method depends on the amount of samples on the cover slide, e.g. on a 60k microarray the sample is hybridized on 60 000 oligonucleotides, corresponding 41 KB overall median probe spacing. As reference DNA we used commercially available male reference DNA (Promega, Madison, WI, USA), which was

amplified for the hybridization with cfDNA. Amplified reference DNA as well as amplified plasma-DNA were labeled with the Bioprime array-CGH genomic labeling system (Invitrogen) and hybridized on slides covered with synthetically produced oligonucleotides. Genomic imbalances can be detected by comparison of the fluorescence intensity between reference and plasma-DNA. Slides were scanned using a microarray scanner, and images were analyzed using Feature Extraction and DNA Workbench 5.0.14 (Agilent Technologies) with the statistical algorithm ADM-2, the sensitivity threshold was 6.0.

For hybridization 2 µl of Cot-1 DNA, 4.5 µl Agilent 10 x Blocking Agent and 22.5 µl Agilent 2 x Hybridization Buffer were mixed with labeled reference- and sample DNA.

The sample was denatured for 3 min on 95 ° C and afterwards incubated at 37°C for 30 min. 45 µl of the hybridization-mix were added on 8 x Gasket-Slide closed on the microarray in a hybridization oven and hybridized for 16 – 24 hours at 65°C.

4.2.10.2. Array-CGH evaluation

(Published in Heitzer *et al.*, 2013).

Evaluation of our array-CGH was done based on our previously published algorithm (Geigl *et al.*, 2009) in R (R Development Core Team 2008). In brief, data normalization and calculation of ratio values were conducted employing the Feature Extraction software 9.1 from Agilent Technologies. The algorithm focuses on detecting which ratio values differ significantly from the ratio profile's mean and should therefore be considered as over- or underrepresented. The concept of the algorithm includes the employment of running means with different window sizes and analyses at progressively greater levels of smoothing and subsequently combining these analyses. Consecutive data points are combined and their mean ratio values are presented in graphs of array-CGH results. The algorithm iterates through the profile by changing window positions, employing a sliding window approach. The algorithm calculates the mean ratio value for each window based on the respective ratio values. Assuming that a window's ratio values are distributed normally, we estimate the standard deviation (SD) by considering the outmost value that is within $\pm 34.1\%$ of the mean. Thresholds were defined as ± 1.25 times the SD in single cell experiments and, due to the higher noise with plasma-DNA, thresholds had to be defined more stringently as ± 1.5 times the SD. The values obtained were assigned to all oligos of the respective window. If a window shows a significantly increased or decreased mean ratio value the mean position of that window will be displayed above or below the

respective region of the ratio profile. We used 7 different window sizes consisting of 10, 25, 50, 100, 250, 500, and 750 adjacent oligos. Depending on the window size it will be labeled with a different color and distance to the X-axis, thus generating a color bar code. As we use sliding windows, the assessment of the copy number status of each oligo is based on 44 different calculations. The final assessment is indicated in a single green or red bar for gained or lost regions, respectively, which is generated if at least 39 (90%) of the 44 repetitive calculations consistently result in the same copy number change. Furthermore, the algorithm generates a table with all localizations of significant calls which allows detailed mapping of each CNV. All ratio profiles shown in the center of the images were calculated using a 500 oligonucleotide window.

4.2.10.3. Plasma DNA array-CGH artefact correction

(Published in Heitzer *et al.*, 2013). We generated random DNA libraries by converting the plasma DNA fragments to PCR-amplifiable OmniPlex Library molecules flanked by universal priming sites for whole genome amplification (WGA). We subjected the WGA products to array-CGH on a 60k microarray platform.

Copy number estimations after WGA may have to deal with potential amplification artifacts, which may result in under- (e.g. allele drop out) or over-representations (e.g. preferential amplifications). Thus, a systematic amplification bias had not been observed with the single cell amplification products as we had previously reported (Geigl *et al.*, 2007; Geigl *et al.*, 2009), we had not employed corrections for single cell analyses. However, with the plasma DNA WGA products we noted amplification biases and, because a large part of this amplification bias was systematic (e.g. correlated with GC-content of the DNA), we were able to account for this bias as follows. Firstly, we mapped the systematic amplification biases in healthy controls (3 females and 3 males) (Supp. Table 4) (illustrated in Supp. Fig. 9a-b). We then introduced corrections for these regions in our algorithm for plasma DNA. In brief, for the regions listed in Supp. Table 4 we increased the number of repetitive calculations, which had to result consistently in the same copy number change from 39 (90%) of the 44 calculations to 41 (93%), requiring that the copy number change is also indicated at least in the 500 or 750 window size calculation. For the representation of the ratio profile in the center of our array-CGH illustrations these regions were adjusted with a correction factor, which depended on the aforementioned calculations. Female and male array-CGH healthy control profiles before

and after these corrections are shown in Supp. Fig. 9. The same corrections were applied in all plasma DNA analyses of the tumor patients.

4.2.11. BEAMing (Beads, Emulsions, Amplification, and Magnetics) Assay

BEAMing involves single-molecule PCRs on magnetic beads in water-in-oil emulsions, allowing the quantitative detection of somatic mutations with high sensitivity and selectivity (Diehl *et al.* 2005; Diehl *et al.* 2006). Our BEAMing analyses were performed by Inostics GmbH (Hamburg, Germany).

4.2.12. Deep sequencing with the 454 GS FLX genome sequencer

(published in Heitzer *et al.*, 2013).

“Picotiter plate pyrosequencing”, a massively parallel sequencing-by-synthesis approach, relies on emulsion PCR-based clonal amplification of a DNA library adapted to micron-sized beads and subsequent pyrosequencing-by-synthesis of each clonally amplified template in a picotiter plate, generating up to 1.000.000 unique clonal sequencing reads per experiment. Sequence variants that represent a fraction of a complex sample can be vastly oversampled, thus enabling statistically meaningful quantification of low-abundance variants.

For amplicon preparation of *KRAS* we used a proof-reading enzyme (Fast Start High Fidelity System, Roche Diagnostics) according to the manufacturer’s instructions with 35 cycles of amplification. Primers were designed resulting in 3 different amplicon lengths (*KRAS* specific length 119 bp, 168 bp, and 323 bp), including the Roche-compatible adaptors A and B with a length of 21 bp, plus 4 bp TCAG key sequence (each read has to start with key sequence) and 10 bp MID (multiple identifier; barcode). Amplicons from patient and control samples (cell-free native and whole genome amplified DNA from plasma) were purified, quantified with PicoGreen, and pooled equimolarly (6-8 samples per pool). Amplicon pools were quality checked on an Agilent Bioanalyzer using Agilent High Sensitivity DNA kit. Emulsion PCR was performed according to emPCR Method Manual – Lib-A MV. In doing so, two emulsion PCRs were performed for bidirectional sequencing (forward and reverse). Microbeads were enriched and deposited onto PicoTiterPlates provided for the 454-FLX instrument (Roche Diagnostics). Massively parallel pyrosequencing was performed according to the manufacturer's protocol. Base calls and quality scores were generated using the GS Run Processor software on a HPC-cluster, and variants were extracted using the GS Amplicon Variant Analysis 2.6 software provided with the platform. Read lengths strongly corresponded to amplicon lengths. The

average coverage was 57045 x for the 119 bp fragment, 12201 x for the 168 bp fragment, and 13540 x for the 323 bp fragment, respectively. Deep sequencing with fragments even longer than 323 bp is challenging because the number of sequencing reads decreases. For the 470 bp fragment length, the number of reads was too low for unequivocal interpretations. The DNA preparations were sequenced with a GS FLX genome sequencer (454 LifeSciences/Roche Diagnostics) according to the manufacturer's instructions.

4.2.12.1. Deep sequencing of DNA fragments after whole-genome amplification

(Published in Heitzer *et al.*, 2013) We subjected the random DNA libraries generated during the WGA process (from the same patients as above) to deep sequencing to estimate whether shifts between the ratios of mutated versus non-mutated DNA fragments had occurred during the amplification process. We did not detect *KRAS* mutations in the WGA-products of the four patients (#7, #11, #15, #16) with low frequency or absent *KRAS* mutation in their native DNA (Supp. Table 2). For patients with a high frequency of *KRAS* mutation (#6, #10, #15, #38) the percentage of mutated DNA fragments was about the same as in the native DNA in patients #10 and #25; however, this percentage was lower in patient #6 and even notably lower in patient #38 (Supp. Table 2). This suggests that WGA may cause a shift towards a higher number of non-mutated DNA fragments in a subset of cases.

4.2.12.2. Statistics

(published in Heitzer *et al.*, 2013)

Statistical analysis was done using Microsoft Excel or R for the Wilcoxon Rank Sum test, which was used for the comparisons of plasma DNA concentrations and CTCs of patients with and without biphasic plasma DNA size distributions. Unpaired two-sided Student's t-tests were used to calculate *P* values for the other data sets.

5. Results

5.1.1. Analysis of patients after sex-mismatched bone marrow transplantation

Cell-free circulating DNA was not only found in cancer patients. During pregnancy a median of 10% of the DNA in the plasma of pregnant woman is of fetal origin (Lun *et al.*, 2008) and also, donor derived DNA has been found in the plasma of recipients following liver and kidney transplantation (Lo *et al.*, 1998).

During the course of our work we have tried to answer the question which amount of DNA in plasma originates from the hematopoietic system and what is the contribution of tumor to plasma DNA concentration. Therefore, we examined plasma from patients after sex mismatch bone marrow transplantation.

In this experimental setup, DNA in hematopoietic cells of the transplant recipient were predominantly or even exclusively of donor origin, whereas the DNA in non-hematopoietic tissues was from the recipient. Because either the donor or the recipient was male, we used a very sensitive, quantitative real time PCR method to evaluate the amount of Y-chromosomal DNA, especially the sex-determining region of the Y-chromosome, in recipient blood plasma (Lui *et al.*, 2003; Partsalis *et al.*, 2006).

As expected, the analysis of the percentage of Y-chromosomal DNA in plasma revealed differences between the group of females receiving bone marrow from male donors and the group of males receiving bone marrow from female donors. The median percentages of Y-chromosomal DNA in the group of female donors to male recipients were 20%, whereas the content of Y-chromosomal DNA in the group of male donors to female recipients was about 80%. In male patients receiving bone marrow transplantations from female donors, the *SRY* sequence detected mainly Y-chromosomal DNA from non-hematopoietic cells of the male recipients. As the Y% of plasma reflected the percentage of free plasma DNA originating from non hematopoietic cells of the body (Lui *et al.*, 2003), about 20% and 80% were derived from non-hematopoietic and hematopoietic cells, respectively. The result of all 20 sex-mismatch bone marrow transplantation patients is shown in the following figure (Fig. 10).



Figure 10: Box blot showing the percentage of Y-chromosomal DNA (Y%) in plasma of sex- mismatch bone marrow transplantation patients. Male recipients with female donors showed about 20% of Y-chromosomal DNA, whereas female recipients with male donors exhibited about 80%.

In cancer patients we assumed that the proportion of DNA from the hematopoietic system is decreased whereas the fraction of DNA from non-hematopoietic cells is increased as the tumor releases additional DNA into the circulation.

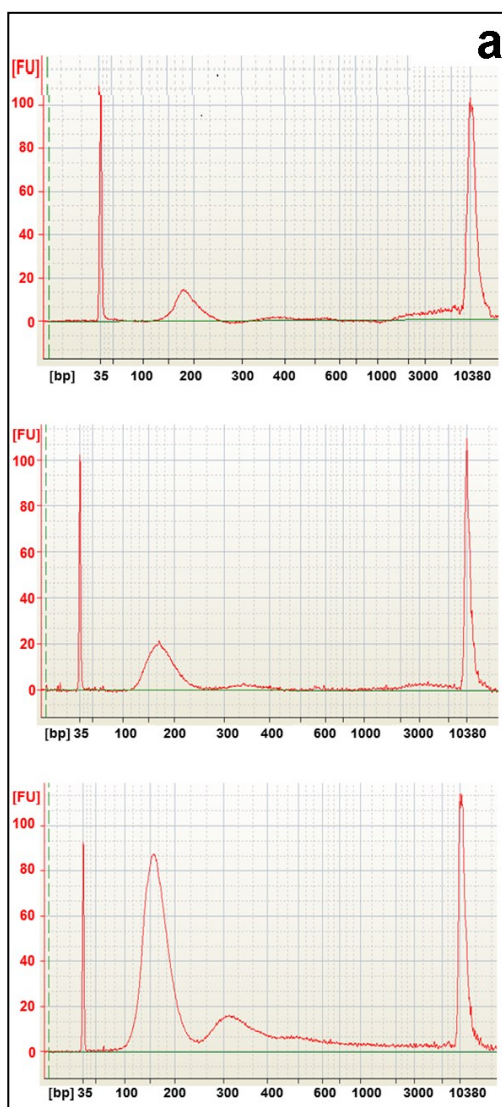
We isolated plasma-DNA of 32 CRC-patients and established protocols for DNA quantification and estimation of DNA fragment size. Furthermore, we developed a protocol for whole genome amplification and compared plasma DNA with DNA from the primary tumor and/or metastasis by array-CGH.

5.1.2. Analysis of patients with CRC

5.1.2.1. A subset of patients with CRC had a biphasic plasma DNA size distribution

Data were published in Heitzer *et al.* (2013):

We analyzed the peripheral blood of 32 patients with advanced stage CRC (Supplementary Information and Supplementary Table 1). We started with measuring the plasma DNA concentrations of healthy controls and patients with advanced-stage CRC. Compared to controls (mean: 15.21 ng/ml; median: 14.37 ng/ml; range: 12.20-19.51 ng/ml), patients showed invariably higher values with substantial variability (mean: 275.35 ng/ml; median: 139.0 ng/ml; range: 22.44-1,037.49 ng/ml) ($P < 0.0001$).



Because mutant DNA fragments in the blood circulation of cancer patients have been reported to be degraded relatively to non-mutant DNA fragments (Diehl *et al.*, 2005), we used a microfluidics-based platform for sizing. We observed an enrichment of plasma DNA fragments within the range of 85-230 bp in the healthy controls (Fig. 11a, upper plot). Plasma DNA fragments within this size range have been associated with the release of DNA from apoptotic cells after enzymatic processing (Diehl *et al.*, 2005; Lo *et al.*, 2010). There was no significant qualitative difference regarding the size distribution of plasma DNA fragments between the healthy controls and 21 (65.6%) of the CRC cases despite the higher concentrations of plasma DNA in the latter group (Fig. 11a, center plot). However, in 11 (34.4%) patients we observed a second peak consisting of DNA fragments with a size range of 240 bp to 400 bp, or even longer in some cases (Fig. 11a, lower plot).

Figure 11a: Characteristics of plasma DNA from healthy controls and patients with advanced-stage colorectal carcinoma (CRC). (a) Size distribution of plasma DNA fragments from a healthy donor (upper plot), patient #11 (center plot) and patient #22 (lower plot). Normalization was performed using two internal markers, visible as high, narrow amplitudes at positions 35 bp and 10.380 bp, respectively. For each analysis, 800 pg of DNA was used.

Interestingly the patients with biphasic plasma size distribution had significantly higher plasma DNA concentrations (mean: 604 ng/ml; median: 562 ng/ml; range: 260-1037 ng/ml) than patients without a second peak (mean: 103 ng/ml; median: 89 ng/ml; range: 22-201 ng/ml) ($P < 0.0001$) (Fig. 11b).

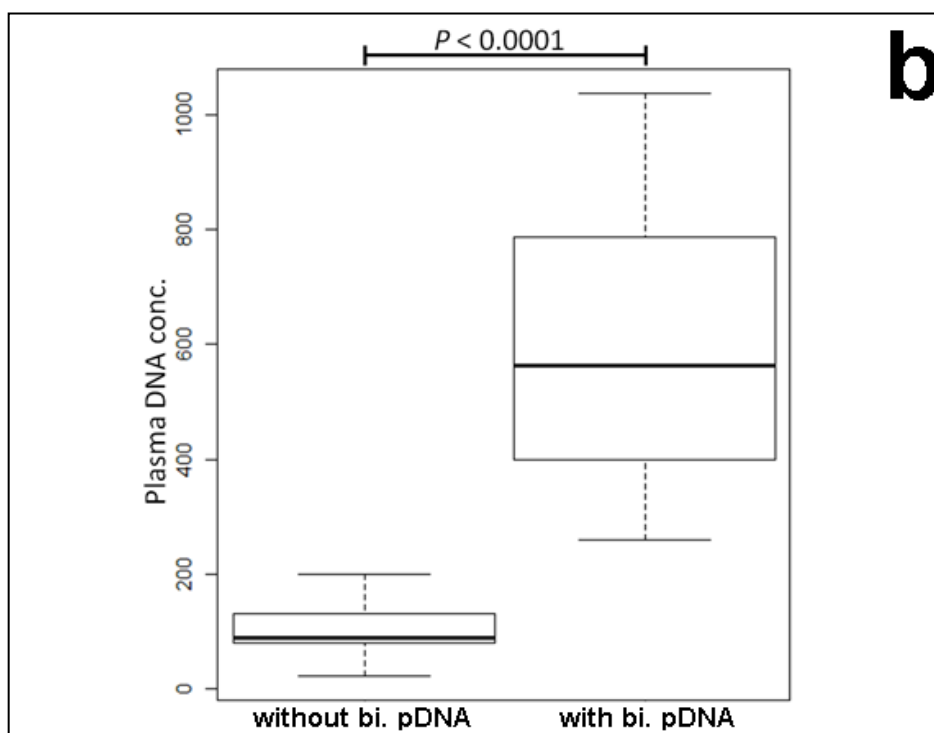


Figure 11b: Patients with a biphasic plasma DNA size distribution (with bi. pDNA) have higher plasma DNA concentrations compared to patients lacking the second peak (without bi. pDNA).

5.1.2.2. The biphasic plasma DNA size distribution correlates with CTC occurrence

To investigate whether the detected DNA size distributions might reflect the contributions of different tumor cell populations releasing their DNA into the circulation around the time of blood collection, we determined the number of CTCs. We used the FDA-approved Veridex system for CTC detection in 30 of the 32 patients (the analysis could not be performed for patients #10 and #35). We observed a clear correlation between the presence of a biphasic plasma DNA size distribution and CTC occurrence. In the patient group with a biphasic distribution (n=10) we found a mean number of 52 CTCs (median: 35; range: 0-181). In contrast, the mean CTC number in patients without a biphasic plasma DNA size distribution (n=20) was only 1.5 (median: 1; range: 0-7) ($P=0.0003$) (Fig. 11c).

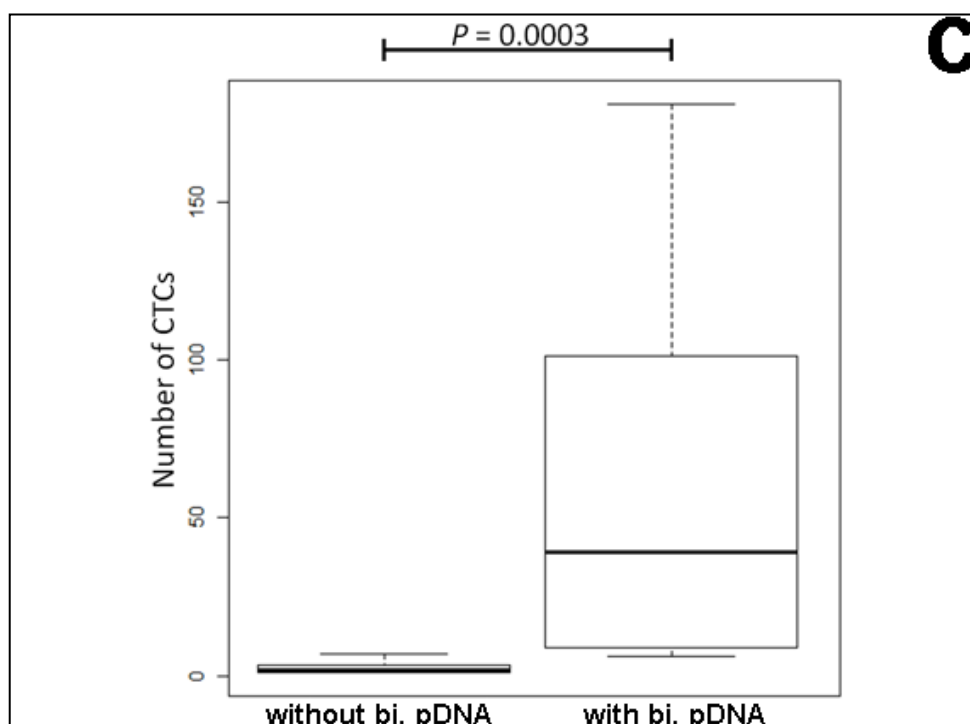


Figure 11c: The occurrence and number of CTCs is closely correlated with a biphasic plasma DNA size distribution.

5.1.2.3. Deep sequencing of plasma DNA for KRAS mutations

To investigate whether mutant DNA fragments were present in both the first and second peaks in the plasma DNA fragment size distribution, we employed ultra-deep pyrosequencing. We used 8 patient samples with and without biphasic plasma DNA size distributions and *KRAS* mutations in their corresponding primary tumors to establish the percentage of mutated plasma DNA fragments for different sequencing reaction sizes (i.e., 119 bp, 168 bp, and 323 bp). In 4 patients (#6, #10, #25, and #38) with biphasic plasma DNA size distributions we found a high percentage of mutated DNA fragments compared to the other 4 patients without biphasic peak (#7, #11, #15, and #16) (Fig. 11d; Supp. Table 2). In these four patients mutated DNA fragments were not observed in the 323 bp sequencing reaction.

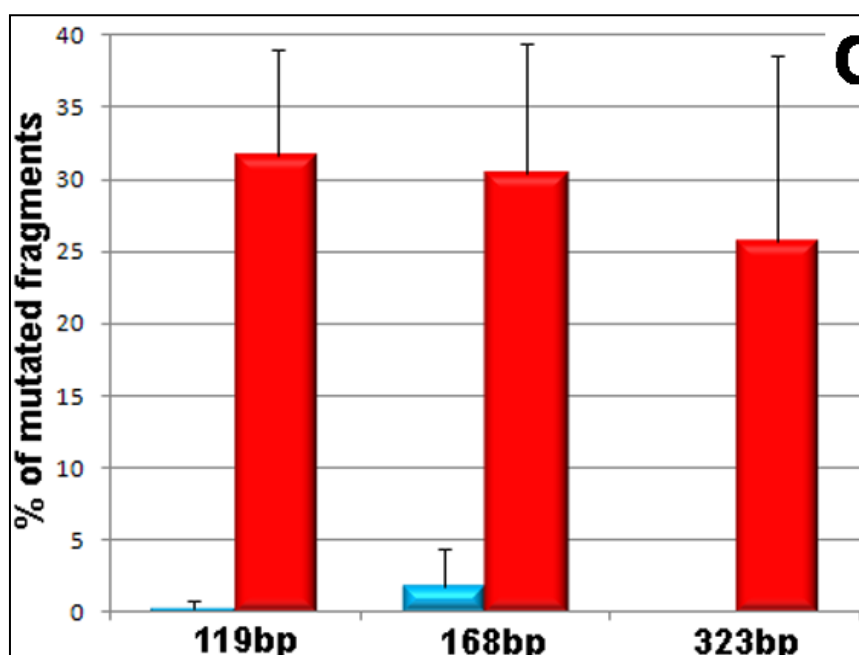


Figure 11d: Deep sequencing using 3 different sequencing reaction sizes (i.e., 119 bp, 168 bp, and 323 bp) identified few mutated *KRAS* fragments in patients without a biphasic plasma DNA size distribution (blue), but high levels of mutated *KRAS* fragments in patients with a biphasic plasma DNA size distribution (red) (the errors bars represent SDs).

5.1.2.4. Genome-wide estimation of copy number changes in plasma DNA

We postulated that, in addition to *KRAS*-mutated DNA fragments, other tumor DNA fragments should be present in the circulation of these patients and that these might provide insights into the tumor genome. To address this hypothesis, we generated random DNA libraries by converting the plasma DNA fragments into PCR-amplifiable OmniPlex Library molecules flanked by universal priming sites for whole-genome amplification (WGA) and subjected the WGA products to array-CGH on a 60 k microarray platform.

This array platform consists of 55,077 oligonucleotides and we calculated for each of these oligonucleotides whether the ratio values were decreased, balanced, or increased. Plasma DNA from 6 healthy controls showed a mean of 4,026 oligonucleotides from the 55,077 oligonucleotides (7.2%; range: 3,606-4,321; 6.4%-7.7%), with aberrant ratio values for the autosomes (Fig. 12a), (please see Supplementary Information for detailed calculations). When we performed array-CGH with plasma DNA from patients who had only the first plasma DNA peak (n=21), we observed an increase in oligonucleotides with aberrant ratio values with a mean of 5,572 oligonucleotides (10.0%; range: 4,587-7,469; 8.2%-13.4%; $P=0.001$ compared with the healthy controls) (Fig. 12b).

Because these oligonucleotides involved only single, non-adjacent oligonucleotides, they most likely represented artifacts of the fragmentation and/or amplification process. In contrast, the plasma DNA in 10 of the 11 patients with a second peak had a mean number of oligonucleotides with copy number changes of 12,476 (22.3%; range: 9,117-14,915; 16.3%-26.7%), which differed highly significantly from both the aforementioned CRC cases and the healthy controls ($P < 0.0001$ each) (Fig. 12c)

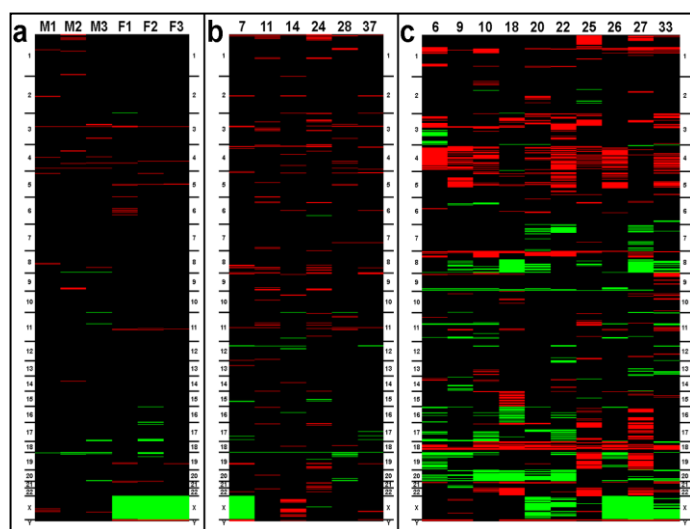


Figure 12: Heat maps of plasma DNA profiles from healthy controls and patients with advanced-stage colorectal carcinoma (CRC).

(a) Heat map of plasma DNA profiles from healthy donors (black: balanced; red: under-represented; green: over-represented). As we used male reference DNA in all experiments, female plasma DNA samples have a relative over-representation of the X chromosome and an under-representation of the Y-chromosome (F1-F3), whereas male samples have balanced sex chromosomes (M1-M3).

(b) Heat maps of six exemplary patients (#7, #11, #14, #24, #28, and #37) that lack a biphasic plasma DNA size distribution.

(c) Heat maps of plasma DNA from 10 patients (#6, #9, #10, #18, #20, #22, #25, #26, #27, and #33) that have a biphasic plasma DNA size distribution.

5.1.2.5. Tumor-specific, genome-wide imbalances in plasma DNA with a biphasic size distribution

Because most of the aforementioned copy number changes in the plasma DNA with biphasic size distribution involved large regions of adjacent oligonucleotides, we determined whether or not they were tumor specific by comparing them with copy number changes of the primary tumor and, if available, metastatic material. As an exemplary case we present patient #6 in greater detail. When the initial diagnosis was made in patient #6 he already had metastases in liver, bones, abdominal lymph nodes, and peritoneum. At this time only a biopsy was obtained from the primary tumor. Five months later a cerebellar metastasis was completely resected, and 3 months after that - i.e. 8 months after the initial diagnosis - we collected blood for our analyses. Thus, multiple tumor sites could have released tumor DNA into the circulation at the time of blood collection.

We noted marked copy number differences between the primary tumor and the cerebellar metastasis in patient #6 (Supp. Fig. 1, panels 1 and 2), indicating the presence of several

malignant cell clones. We found numerous copy number changes in the plasma DNA (Supp. Fig. 1, panel 3) and performed a detailed analysis consisting of the following steps: First, we compared the ratio of the profiles of the available material, i.e. primary colorectal tumor, cerebellar metastasis, and plasma DNA (Fig. 13a).

We then constructed detailed heat maps showing whether the copy number status was decreased, balanced, or increased for each oligonucleotide on our array platform (Fig. 13b). Finally, we determined whether the copy number status for each oligonucleotide occurred only in the plasma DNA or also in the primary tumor and/or metastasis (Fig. 13c).

These analyses revealed that the copy number status of 46.3% of the oligonucleotides on our array platform was present in all three lesions; 18.1% was shared by the primary tumor and plasma DNA, 18.0% was shared by the metastasis and plasma DNA, and 17.6% was unique to the plasma DNA (Fig. 13c).

This suggested that the observed changes in the plasma DNA were tumor-specific and were largely caused by clones both from the primary and the cerebellar metastatic tumor. As the cerebellar metastasis had been removed, cells from this metastatic clone were apparently still present in the patient. Whether the copy number changes observed only in the plasma DNA could have resulted from the other metastases or from parts of the primary tumor that were not included in our analysis remains unknown.

In addition to patient #6 we had 3 other patients (i.e. #9: Supp. Fig. 2; #26: Supp. Fig. 3; #33: Supp. Fig. 4) for whom material from the primary tumor and metastases was available for comparison. Altogether, the plasma DNA samples from these 4 patients (i.e. #6, #9, #26, and #33) displayed an identical copy number status (i.e. lost, balanced, gained) in all three samples (primary, metastasis, and plasma DNA) for an average of 54.7% (median: 54.3%; range: 46.3-63.6%) of all oligonucleotides.

About 10% (median: 9.4%; range: 3.1-18.1%) of plasma DNA copy number changes were unique to the metastasis and, vice versa, approximately 14.5% (median: 16.2%; range: 2.9-22.7%) were only present in the primary tumor. About 20.9% (median: 18.7%; range: 15.7-30.4%) of plasma DNA changes were observed only in the plasma DNA but not in the primary tumor and metastasis. However, all 4 patients had additional metastases at various sites (Supp. Table 1), which were not accessible to us.

Thus, copy number changes observed only in the plasma DNA could reflect alterations from a metastatic site not included in our analysis.

For another patient (#27) we had material from the primary tumor only and we obtained comparable results with the plasma DNA, i.e. copy number changes similar to those observed in the primary tumor (Supp. Fig. 5).

By contrast, we did not find copy number changes of large contiguous chromosomal regions in the plasma DNA of patient #38 that unequivocally correspond to those of the primary tumor or metastases (Supp. Fig. 6).

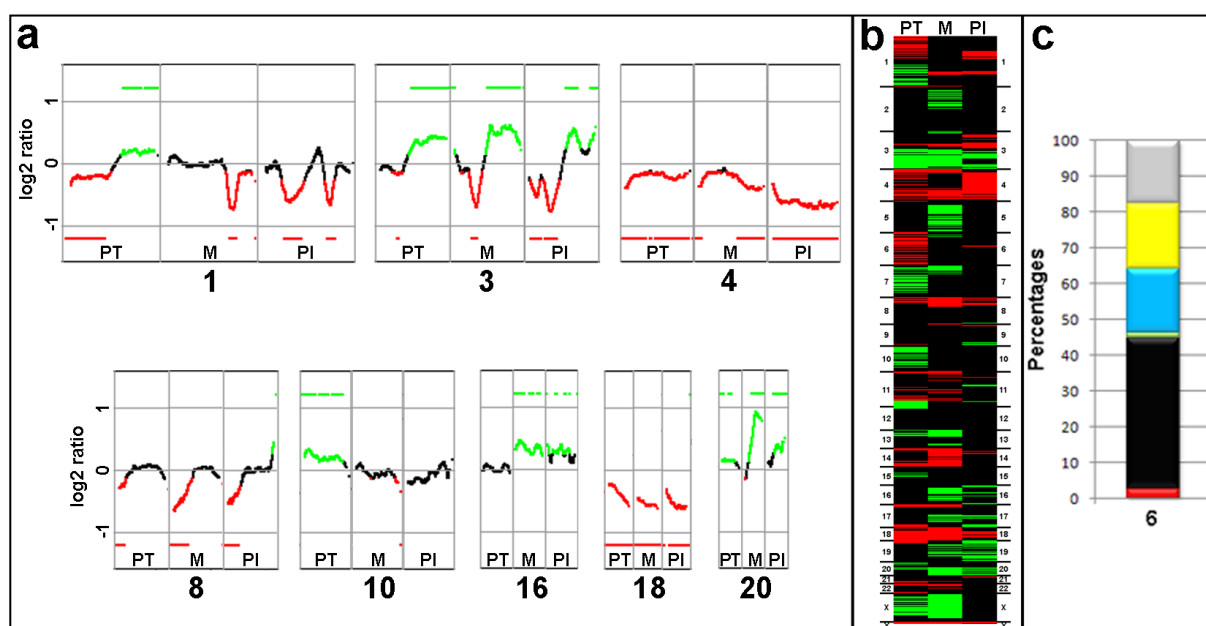


Figure 13: Comparison of copy number changes in the primary tumor, metastasis and plasma DNA of patient #6.

(a) Comparison of the ratio profiles of representative chromosomes (1, 3, 4, 8, 10, 16, 18, and 20; indicated by numbers below the copy number changes) between the primary tumor (PT, left column), metastasis (M, center column), and plasma DNA (PI, right column). The single green and red bars summarize the regions that were gained or lost based on all iterative calculations of our algorithm (Supplementary Information). The black profile regions represent balanced regions, lost regions appear in red, and gained regions are shown in green (complete profiles are depicted in Supplementary Fig. 1a).

(b) Heat maps comparing the copy number changes in the primary tumor (PT), metastasis (M), and plasma DNA (PI); Black: balanced regions; red: under-represented regions; green: over-represented regions).

(c) The bar chart combines information on copy number changes and their occurrence in the primary tumor, metastasis, and plasma DNA. It displays the percentages of chromosomal regions that were commonly lost (red), balanced (black), or gained (green) in all three samples, shared by metastasis and plasma DNA only (blue), shared by primary tumor and plasma DNA only (yellow), or unique to the plasma DNA (gray).

5.1.2.6. Plasma DNA analysis in the absence of material from the primary tumor or metastasis

In three cases, i.e. #10 (Supp. Fig. 7a), #20 (Supp. Fig. 7b), and #25 (Supp. Fig. 7c), no further tumor material was available for analyses. However, in each of these cases we were

able to establish array-CGH profiles reminiscent of copy number changes frequently observed in colon cancer (<http://www.progenetix.net>), such as loss on 8p and gains on 8q and 20.

In two cases (i.e. #18, #22) insufficient material was available, as only small biopsies had been taken at the time of diagnosis. In patient #18 our plasma DNA analysis again revealed typical CRC-related copy number changes, such as loss on 8p and gains on 8q and 20 (Fig. 14a), whereas in patient #22 there were losses on chromosomes 3, 4, 5, 8p, and 18 and gains on chromosomes 7p, 17q, and 20 (Fig. 14b). These results suggest that tumor-specific copy number profiles can be established even in cases where no material from the primary tumor or metastasis is available for various reasons.

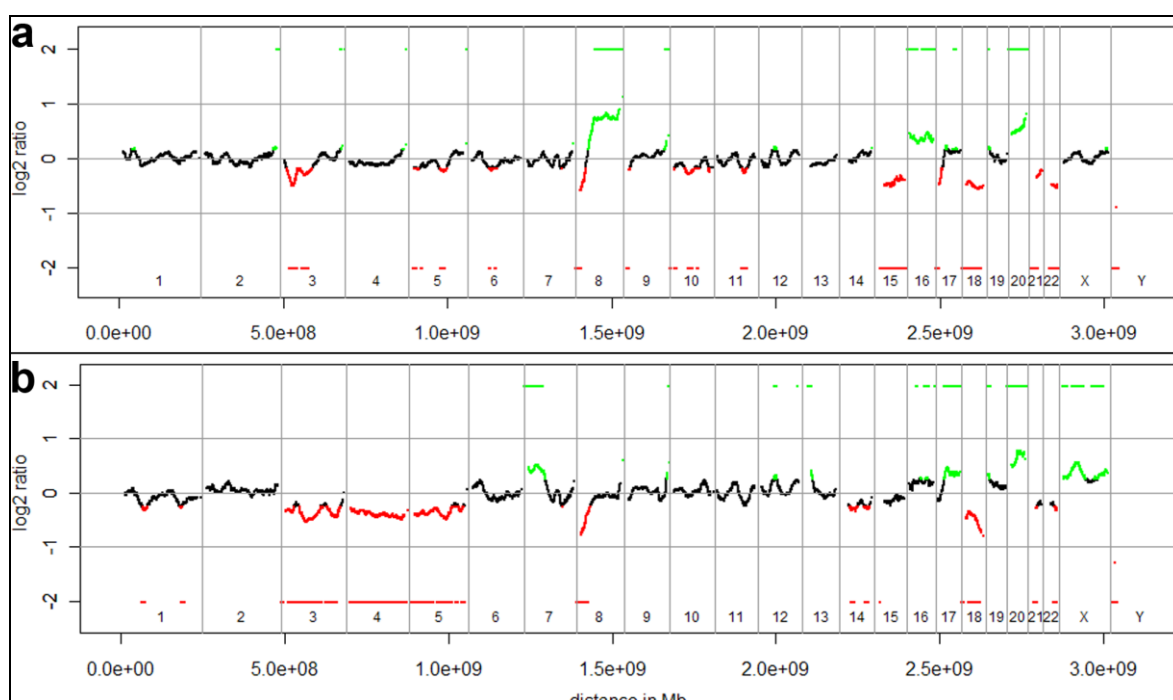


Figure 14: Array-CGH from two cases where only small biopsies had been taken at the time of diagnosis, so that insufficient material was available for further analyses.

(a) The plasma DNA from patient #18 revealed typical CRC-related copy number changes, such as loss on 8p and gains on 8q and 20.

(b) Plasma DNA ratio profile from patient #22 demonstrating losses on chromosomes 3, 4, 5, 8p, and 18 and gains on chromosomes 7p, 17q and 20.

5.1.2.7. BEAMing analysis

As all patients included in this study had metastatic disease showing disease progression at the time of blood collection, the significant variability of plasma DNA size distribution and concentrations, as well as CTC number was unexpected. Furthermore, we were puzzled that even deep sequencing did not find evidence for overt mutant DNA fragments in some of our patients. Therefore we tested patients for the presence of *KRAS* mutations with a more sensitive approach, i.e. BEAMing, which has the capacity to detect and enumerate mutant and wild-type DNA when present at ratios greater than 1:10,000 (Diehl *et al.*, 2005; Diehl *et al.*, 2006) (www.inostics.com). To exclude any variation from sample collection and handling we performed serial analyses with blood samples from two patients (L1 and L2) collected on five (L1), respectively four (L2) consecutive days. These patients had end-stage, highly metastatic colon cancer and received only palliative treatment (no chemotherapy or radiation therapy), which should have no impact on the tumor burden. In both patients the *KRAS* G12V mutation had been detected in the respective primary tumors.

In patient L1 the respective mutant fractions were close to the BEAMing detection limit, given as 0.02, on days 1, 2, 4, and 5, or even below the detection limit on day 3 (Fig. 23a). Accordingly, we observed invariably balanced array-CGH profiles on each day. This suggests that even ultrasensitive methods may be incapable of detecting mutant DNA fragments in some patients with highly metastatic disease.

In contrast, in patient L2 the mutant fractions were between 6.644 (day 2) and 11.181 (day 3) (Fig. 15a). On day 3 there was a biphasic plasma DNA size distribution and in the array-CGH analyses we observed copy number changes on chromosomes 3, 4, 18, 19, and 20, which we had previously also observed in the primary tumor (Fig. 15b). Although the plasma DNA size distribution was monophasic on the other days, some of these copy number changes were also visible in the array-CGH profiles on these days.

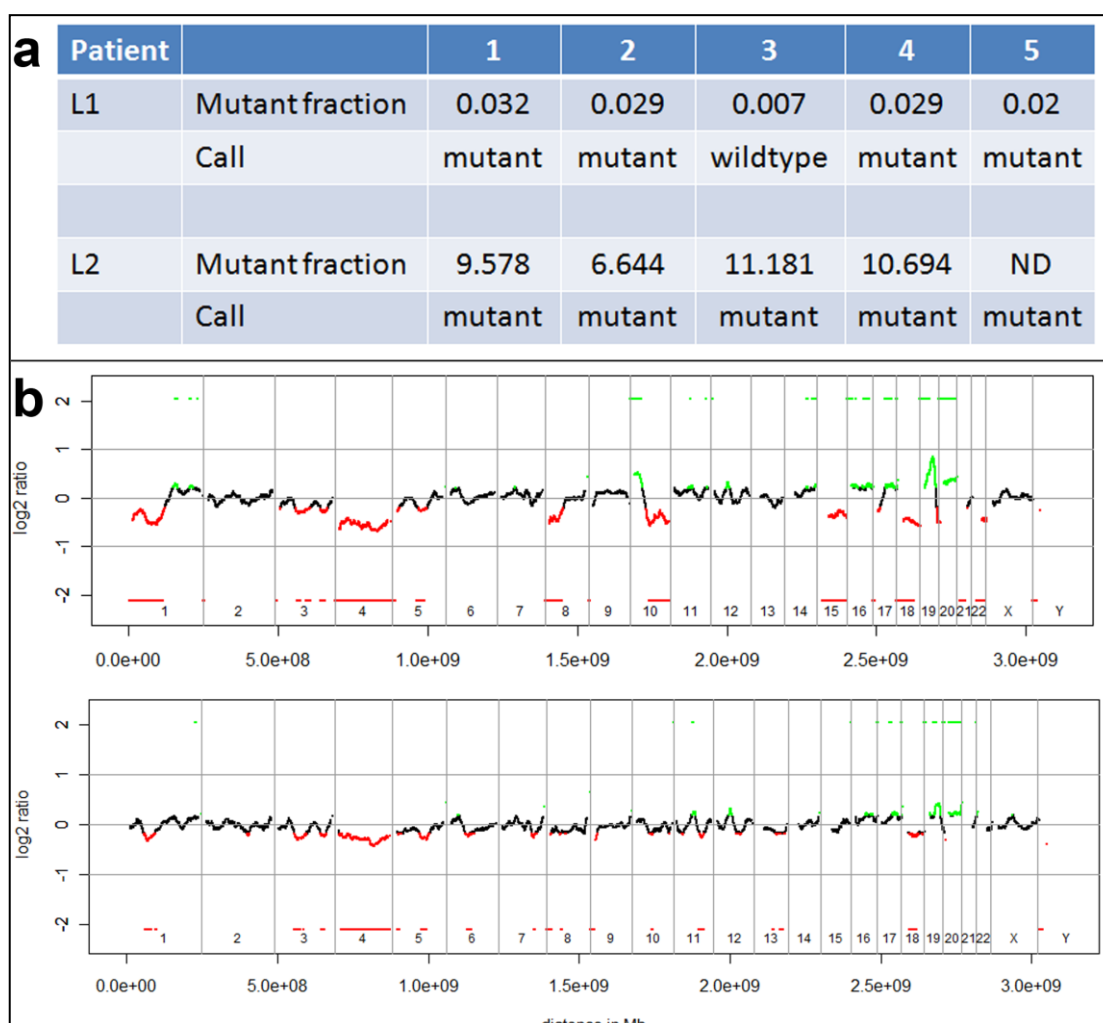


Figure 15: Summary of the BEAMing analysis and representative array-CGH profiles from one representative patient.

(a) Mutant fraction and call (i.e. either mutant or wildtype) as established by BEAMing for two patients.

(b) The upper panel illustrates the array-CGH profile of the primary tumor, the lower panel the corresponding profile established with plasma DNA from day 3.

5.1.2.8. Correlation with clinical parameters

Although this was a pilot study, we then attempted to correlate our findings on plasma DNA size distribution and CTC number with clinical parameters. We did not find any correlation between the levels of the established CRC tumor markers CEA (Supp. Fig. 8a) and CA19-9 (Supp. Fig. 8b) and plasma DNA concentration or CTC numbers. The percentage of patients with a biphasic plasma DNA size distribution was 67%, 44%, and 38% for patients with metastases in bone (n=6), liver (n=25), and peritoneal carcinomatosis (n=8), respectively, and only 9% for patients with lung metastasis (n=11) (Supp. Info.; Supp. Table 3). These percentages were almost identical for patients with more than 6 CTCs, i.e., 67%, 36%, 9%,

and 38%, for patients with metastasis in bone, liver, lung, and peritoneal carcinomatosis, respectively. In fact, only one of the 11 patients with lung metastasis (patient #38) demonstrated both a biphasic plasma DNA size distribution and a high CTC number, according to the Veridex system.

5.1.2.9. Analyses of plasma DNA from patients with breast cancer

Analysis of breast cancer samples was performed by Martina Auer.

To test whether our observations can be extended to other tumor entities we collected blood from 35 patients with breast cancer. Indeed, we observed the biphasic plasma DNA size distribution in 14 (40%) patients, which again correlated with the plasma DNA concentration (Fig. 16a) and number of CTCs (Fig. 16b). In these cases we could again reconstruct tumor specific copy number changes from the plasma DNA (Fig. 16c-d).

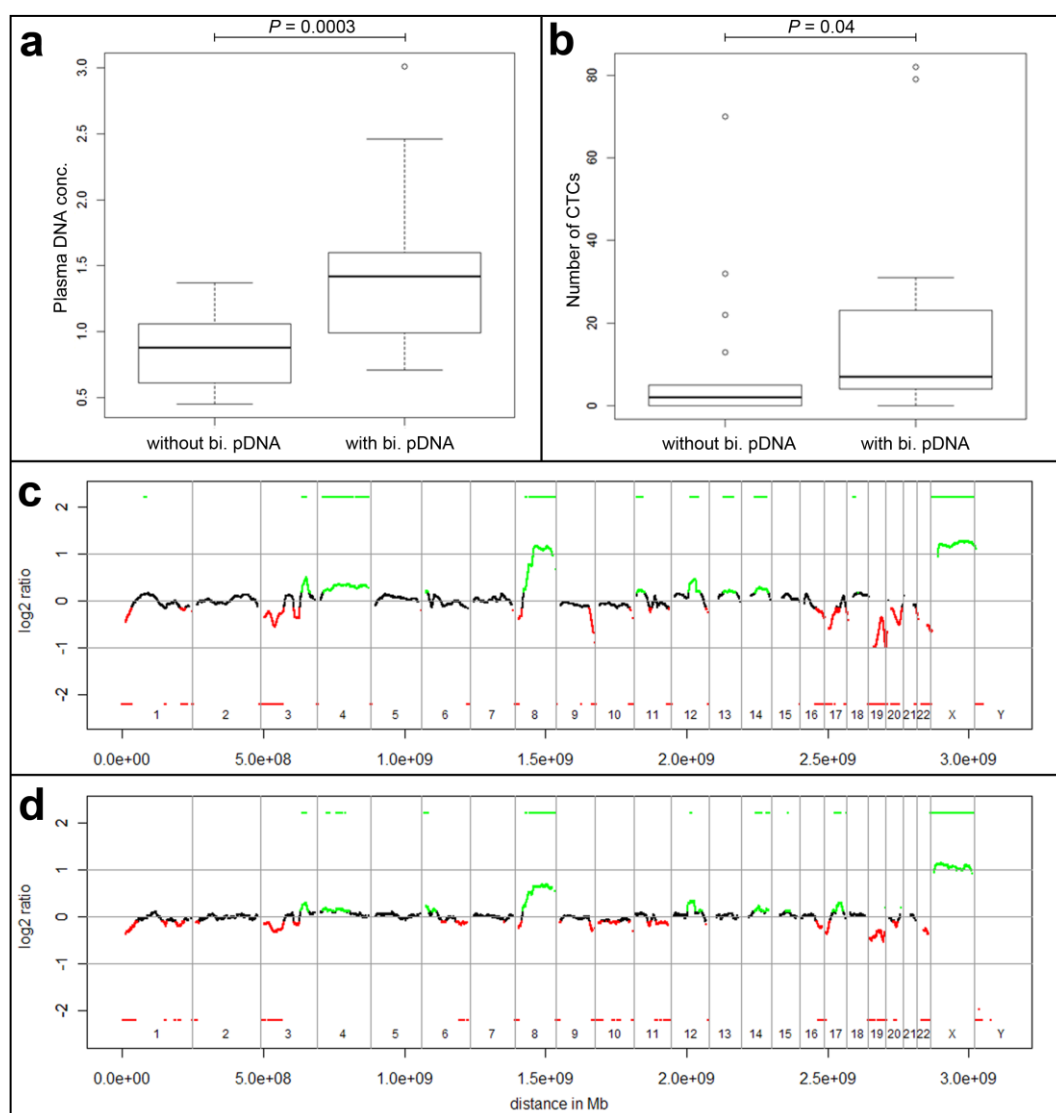


Figure 16: Evaluation of blood from 35 patients with breast cancer and representative array-CGH profiles.

(a) Breast cancer patients with a biphasic plasma DNA size distribution (with bi. pDNA) have increased plasma DNA concentrations compared to patients lacking the second peak (without bi. pDNA). (b) The number of CTCs is also increased in patients with a biphasic plasma DNA size distribution. (c) Array-CGH profile of the primary tumor of breast cancer patient #20.

6. Discussion

Our initial experiments on plasma DNA, performed on peripheral blood obtained from patients after sex-mismatch bone-marrow transplantation, demonstrated that highly fragmented DNA in the blood circulation is detectable. We established protocols for DNA amplification and found methods for quantitative and qualitative analysis, which have been applied in processing the samples from CRC patients.

A major goal of cancer medicine is to move from fixed treatment regimens to therapies tailored to a patient's individual tumor. Following recent breakthroughs in genomics technology, major efforts to identify potentially informative mutations using next-generation sequencing are under way (Roychowdhury *et al.*, 2011). However, such detailed analyses are usually performed with samples obtained during the initial diagnosis and/or metastasis (Bignell *et al.* 2010; Kan *et al.* 2010). All of these genomic studies have confirmed that tumor genomes are complex and highly prone to changes. In this study we addressed whether complex tumor genomes may be inferred non-invasively from the peripheral blood of patients with cancer.

We observed that the plasma DNA and CTCs varied significantly in our cohort, although all patients included in this study had metastatic disease showing disease progression at the time of blood collection. Some cancer patients with only the first peak have higher plasma DNA levels compared to healthy controls, but a very low percentage of mutated DNA fragments, as confirmed by deep sequencing, BEAMing and array-CGH. This is consistent with necrotic neoplastic cells being engulfed by macrophages, which involves the killing of neoplastic cells and the surrounding stromal and inflammatory cells (Diehl *et al.*, 2005). The released DNA contains multiple wild-type DNA sequences, which may explain the increase in total, non-mutant circulating DNA.

In contrast, a biphasic plasma DNA size distribution may indicate a distinct biological process, because its occurrence is associated with very high plasma DNA levels, elevated percentages of mutated DNA fragments in the circulation, and an increased number of CTCs. The biphasic plasma DNA distribution likely reflects massive cell destruction with direct shedding of DNA from tumor cells and cellular fragments into the bloodstream. As the number of CTCs identified by the Veridex system is relatively small, CTCs by themselves likely contribute only small amounts of tumor DNA into the circulation. Hence, the biphasic plasma DNA size distribution more likely results from various mechanism of release at the

site of the tumor, for example when the tumor has invaded through blood vessels. The first peak consisting of plasma DNA fragments within the range of 85-230 bp could be released from apoptotic cells after enzymatic processing. Previously, plasma DNA fragments within this size range had been associated with the release of DNA from apoptotic cells after enzymatic processing, because the length of these fragments corresponds approximately to the DNA wrapped around a nucleosome (~142 bp) plus a linker fragment (~20 bp) (Diehl *et al.*, 2005; Lo *et al.*, 2010). Accordingly, the DNA fragments from the second peak likely represent di- and trinucleosomal DNA fragments. Hence, the co-occurrence of larger DNA fragments suggests that their broken nuclear products were subject to shorter degradation times in the circulation. Alternatively, the biphasic DNA may be a consequence of saturation of DNA degradation mechanisms in the plasma due to high amounts of DNA. In contrast, we did not observe significant amounts of DNA fragments from necrotic cells which were associated with DNA fragment sizes larger than ~10,000 bp (Jahr *et al.*, 2001). Thus, our results may explain previously reported differences in the detection rates of abnormal forms or quantities of DNA in plasma (Schwarzenbach *et al.*, 2011; Pinzani *et al.*, 2010) and why most studies found that plasma tumor DNA fragments are degraded (Muller *et al.*, 2008; Gormally *et al.*, 2007; Diehl *et al.*, 2008b; Forsheew *et al.*, 2012), whereas only few reported an increased integrity (Pinzani *et al.*, 2010; Wang *et al.*, 2003). At present, we do not know whether the biphasic plasma DNA size distribution represents a specific tumor signature, as we have not studied other conditions yet in which plasma DNA may be increased, such as chronic inflammatory diseases.

Our genome wide analyses contrast with numerous previous studies that have attempted to identify abnormal forms or quantities of DNA in plasma or serum (Pinzani *et al.*, 2010; van der Vaart & Pretorius, 2010; Sidransky, 2002). Whereas these previous approaches had usually only monitored a single target from plasma DNA (Diehl *et al.*, 2005; Leary *et al.*, 2010; McBride *et al.*, 2010), we have developed a novel technology to establish genome-wide tumor-specific copy number changes from plasma DNA with easy means. This allows novel diagnostic strategies in cases where sufficient material from the primary tumor is not available for further analyses and it enlarges options for disease monitoring. Recently, Chan *et al.* analyzed plasma DNA from 4 patients with hepatocellular carcinoma and one patient with synchronous breast and ovarian cancer by shotgun massively parallel sequencing to identify tumor-associated copy number changes (Chan *et al.*, 2012). Another study also used a whole-genome sequencing approach for the analysis of plasma of 10 colorectal and breast cancer

patients (Leary *et al.*, 2012). Their results suggest that next-generation sequencing based approaches may increase the resolution of copy number analyses from circulating DNA.

However, our study differs from previous plasma DNA related papers in one important issue. Most studies reported that plasma DNA reflects well the clinical course of a disease and that the amount of mutated DNA fragments would correlate with the tumor burden (Diehl *et al.*, 2008a; Diehl *et al.*, 2008b; Misale *et al.*, 2012; Forshew *et al.*, 2012; Chan *et al.*, 2012). In contrast to these reports, a surprising finding was that some of our patients with progressive, metastatic disease appear to release no mutant DNA fragments in measurable quantities. This was especially exemplified in patient L1, who had highly metastatic, progressive end-stage disease, where even an ultrasensitive approach such as BEAMing was not capable of detecting mutated DNA fragments. In addition, we had several other patients in our cohort of metastatic patients with progressive disease where deep sequencing did not detect mutated DNA fragments. If a close correlation between tumor burden and tumor-specific DNA fragments in the circulation existed, we should have seen tremendous amounts of these fragments in the blood of all our patients. A possible explanation may be that due to the short half-life of plasma DNA, which had been estimated to be merely 16 min (Lo *et al.*, 1999), in some of our cases no significant amounts of tumor DNA had been released into the circulation several hours prior to our blood collection. Hence, testing of the size distribution of plasma DNA fragments may be an easy screening tool to identify blood samples which are worth more detailed plasma DNA and/or CTC analysis.

The next challenge will be to test such approaches for the serial monitoring of cancer patients and to evaluate their potential for the early-stage analysis of peripheral blood. However, our study suggests that non-invasive access to tumor material immediately before specific anti-tumor treatment initiation in advanced cases may become possible for the re-evaluation of the cancer genome and possibly the tailored management of treatment choices.

Part II: DEVELOPMENT OF AN ADENOVIRAL VECTOR FOR THE DETECTION OF CIRCULATING TUMOR CELLS

7. DEVELOPMENT OF AN ADENOVIRAL VECTOR FOR DETECTING TELOMERASE POSITIVE CIRCULATING TUMOR CELLS

7.1. Telomere structure and function

In 1938 Müller described the structures at the chromosomal ends for the first time and named them “Telomeres” according to the Greek words “telos” for “end” and “meros” for “part”. Independently from his work McClintock reported in 1941 that chromosomes with broken ends were able to form dicentric- and ring chromosomes (McClintock *et al.*, 1938; McClintock *et al.*, 1941). In human somatic cells the structural end of the chromosomes are evolutionary highly conserved and characterized by 500 – 3000 tandem repeats of GT-rich sequences, TTAGGG (Fig. 17) (Greider *et al.*, 1998; Moyzis *et al.*, 1988). Also the telomeric orientation is highly conserved – they always run in 5’ to 3’ direction and build the molecular 3’ end of the chromosome (Greider, 1996).

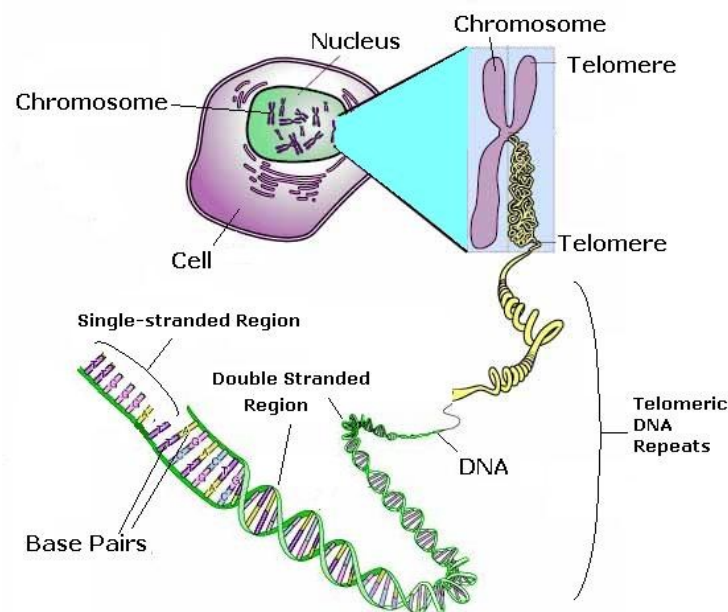


Figure 17: Telomere structure

(Source: <http://www.ch.ic.ac.uk/local/projects/burgoine/origins.txt.html>)

Telomere-associated proteins have been demonstrated to regulate the telomere length. TRF1 induces the bending, looping, and pairing of duplex DNA, whereas TRF2 prevents the fusion of chromosomal ends by formation of a terminal loop structure named the T loop (Fujimoto *et al.*, 2003; Griffith *et al.*, 1999). Further on, telomere-binding proteins are sequence specific, i.e. they protect telomeres against enzymatic or chemical modifications (Blackburn, 1991) and from DNA damage or degradation (Hahn, 2003). Alteration of telomere-binding proteins could lead to genomic instability and influence cell proliferation and/ or cell viability (Blasco & Hahn, 2003).

During the process of semiconservative DNA replication, leading and lagging strands become synthesized. Since the replication fork moves from 5' - 3' the leading strand synthesis is complete. In contrast, synthesis of the lagging strand is performed in the opposite direction to the replication fork. For this purpose short pieces are synthesized in 5' - 3' and ligated to build an integrated fragment. (Ohki *et al.*, 2001; Greider, 1996). Because of the “end replication problem” (Fig. 18), the RNA primer responsible for initiating “lagging strand” synthesis may bind at a site many nucleotides away from the 3' end of the parental strand. As a consequence the ends of linear chromosomes are not synthesized completely during replication and, therefore telomeres lose 50 – 200 bp with every cell division.

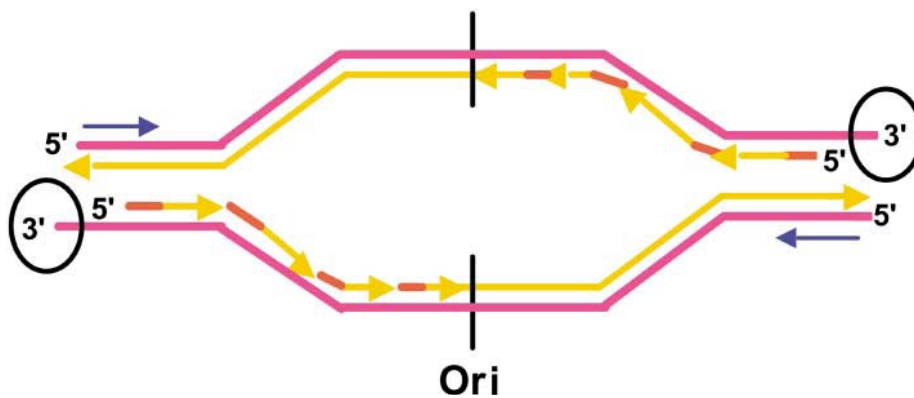


Figure 18: The end replication problem (Source: Campisi *et al.*, 2001). With every cycle of DNA replication 50 – 200 bp of DNA are not synthesized at each chromosomal 3' end. The digestion of replicated telomeres by exonucleases results in 3' overhangs at telomere ends.

In 1961 Hayflick and Moorhead reported a limited replication capacity in normal human fibroblast (Hayflick & Moorhead, 1961). They showed that human fibroblast - from different embryonic donors - were unable to undergo more than 40 – 60 population doublings (PDs), the so called Hayflick limit (Hayflick, 2000). The decrease in telomere length is responsible for the finite number of PD and act as “mitotic clock” (Harley *et al.*, 1990).

Due to the “end replication problem”, telomeres become shorter during each cell division (Harley *et al.*, 1990; Makarov *et al.*, 1997). Consequently, normal cells undergo cellular senescence (cells stop to proliferate, but they are still alive), as telomeres shorten to an average length of 4 – 7 kb (Campisi *et al.*, 2001). One chance to circumvent senescence is the inactivation of p53/Rb tumor suppressor pathways. Then cells will start to proliferate until their telomeres become extremely short. This critical and unstable state is called crisis (Kim *et al.*, 2002; Shay & Roninson, 2004) and it is characterized by karyotypic instability (Sedivy *et al.*, 1998), leading to massive cell death (Maser *et al.*, 2002).

The activation of an enzyme named telomerase or mechanism for the alternative lengthening of telomeres (ALT) ensures the stabilization and lengthening of the telomeres and helps the cell to circumvent crisis (Campisi *et al.*, 2001; Kim *et al.*, 2002; Levy *et al.*, 1992).

7.2. Telomerase

Telomerase is a large ribonucleoprotein complex, responsible for *de novo* addition of TTAGGG-repeats onto the 3' end of the chromosome by using its intrinsic RNA moiety (Fig. 19) (Chan & Blackburn, 2004; Shay & Wright, 2005; Donate & Blasco, 2011). The enzyme consists of two major components: a reverse transcriptase catalytic subunit, hTERT, which is highly conserved and associated with the second subunit, the template RNA, TERC (also referred as TER or TR) (Smogorzewska & Lange, 2004). TERC is used as a template for the synthesis of telomeric DNA (Shay & Wright, 2005).

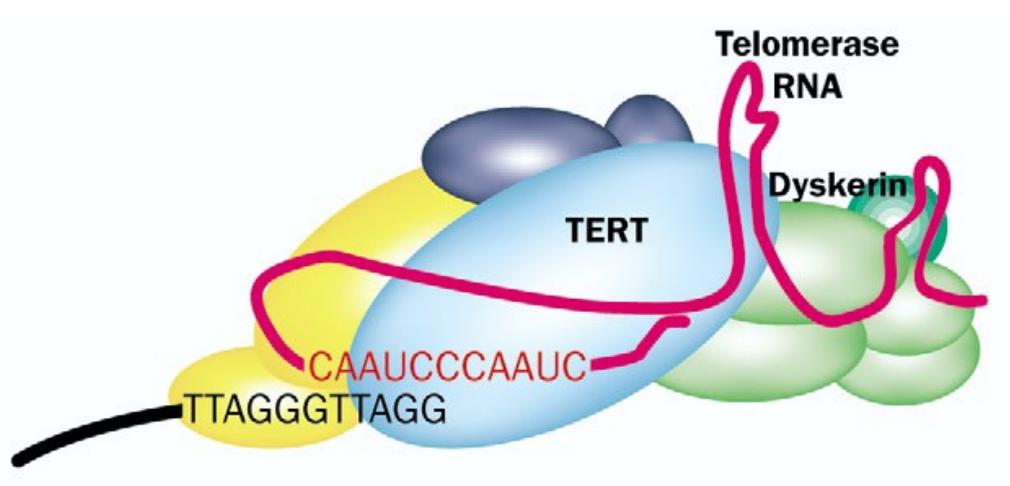


Figure 19: Telomerase ribonucleoprotein complex (Source: Wong & Collin, 2003). Interaction of Telomerase RNA (red) with proteins like Dyskerin (green) leads to an accumulation of stable processed RNA in cells. A stable but inactive version of the ribonucleoprotein complex is present in the majority of somatic human cells. For catalytic activation TERT (blue) has to join the complex. Additional subunits are needed regarding the interaction of telomerase to the 3' end of the chromosome.

Additionally, telomerase is associated with several accessory proteins like dyskerin, which is responsible for telomerase stability and is involved in its biogenesis (Donate & Blasco, 2011). After the ribonucleoprotein complex has synthesized six base pairs, it pauses and translocates the template RNA in order to synthesize the next six base pairs (Fig. 20) (Shay & Wright, 2005). However, telomerase requires nucleotide-triphosphates and primers in 5' - 3' orientation to synthesize new DNA.

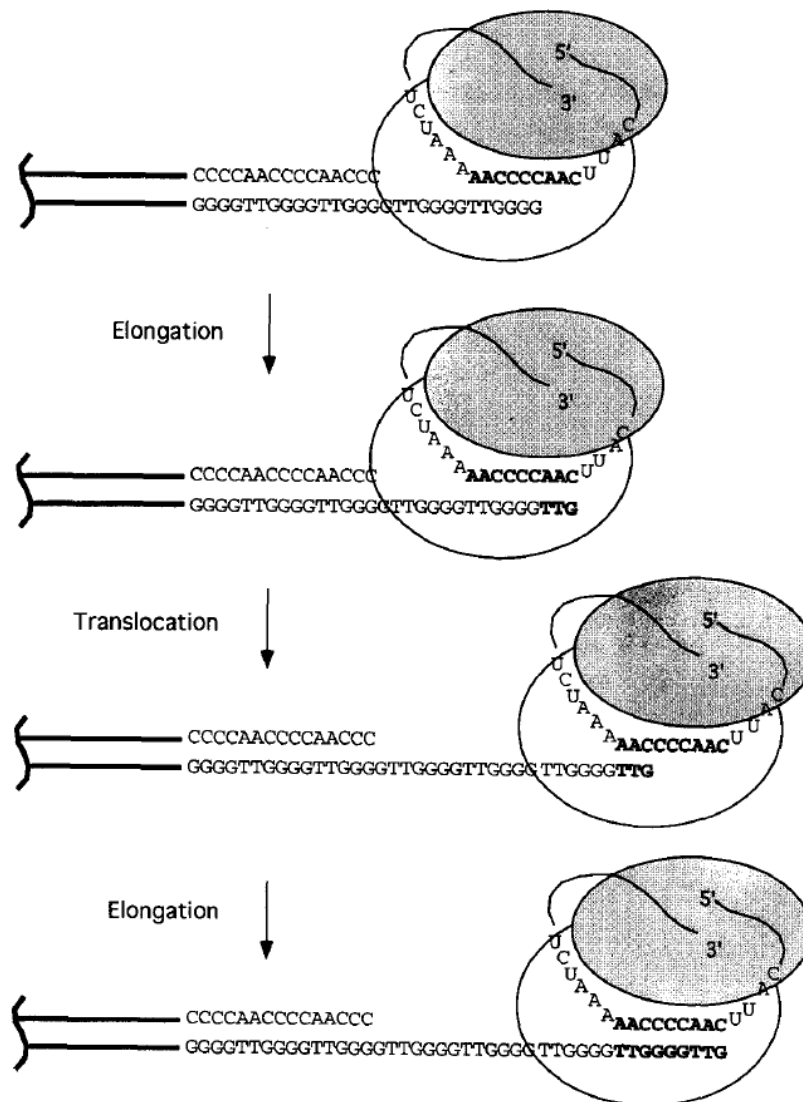


Figure 20: Telomerase lengthening model (Source: Greider, 1991): Telomerase identifies the telomere substrate and the CAACCCAA sequence in the RNA component is paired with the terminal TTGGG repeat. In the next step the copy of the RNA sequence is fused to the template region. Through reposition/ translocation of the TTGGGGTTG sequence an additional template sequence becomes uncovered. By producing additional TTGGG the telomerase elongation process is an evolutionary process to overcome the “end-replication-problem”

7.3. Telomerase activity

Telomerase remains detectable in adult germline cells, haematopoietic stemcell-like cells and epithelial cells (Wong *et al.*, 2003), but also in highly proliferating cells of renewal tissue, like activated lymphocytes (Shay & Wright, 2005). In contrast, in normal somatic cells progressive shortening of telomeres is observed due to the lack of telomerase activity (Shay & Wright, 2005; Bekaert *et al.*, 2004). In 2003 Masutomi and colleagues reported an increased telomerase activity in early passage fibroblasts during cell division. In this context the elevated activity level was not responsible for maintaining telomere length; instead it had an effect on stabilizing the 3'-telomere cap structure (Masutomi *et al.*, 2003).

Besides its important function for cell survival, telomerase and the expression of hTERT protein is associated with an enhanced stability of the genome and an increase in the DNA repair function (Sharma *et al.*, 2003).

In most tumor cells and immortalized cells an elevated telomerase activity compared to normal cells has been reported (Fujimoto *et al.*, 2003; Blasco & Hahn, 2003). Nevertheless, 10 – 15% of tumor cells do not express telomerase. These cells are characterized by extremely long and heterogeneous telomeres, and telomere stability is achieved by a mechanism named ALT (alternative lengthening of telomeres) (Blasco & Hahn, 2003).

7.4. Telomerase promoter

The regulation of the human telomerase activity is organized by transcriptional regulation of the hTERT gene and mechanisms such as alternative splicing or post-translational modifications (Ulaner *et al.*, 1998). Additionally, the up-regulation of hTERT-mRNA-expression displays the critical step in telomerase activation (Takakura *et al.*, 1999).

The hTERT core promoter region, with a length of 181 bp, is adequate to regulate the transcriptional activation in tumor cells (Fig. 21) (Takakura *et al.*, 1999; Horikawa *et al.*, 1999).

In 1999 the telomerase promoter was sequenced and characterized. In several studies the hTERT promoter is described as TATA- or CAAT-box less and CG-rich region, harboring an initiator element nearby the transcription start point (Takakura *et al.*, 1999; Horikawa *et al.*, 1999). Furthermore, the core promoter region comprises two E-boxes serving as binding sites for *Myc*-related family members and, in addition, five binding sites for SP1 (Fig. 30) (Takakura *et al.*, 1999). *c-MYC* plays an important role in the induction of telomerase activity in proliferating and transformed cells and is also important for preserving the chromosomal integrity by maintaining telomere length. *c-MYC* induces the expression of its catalytic subunit TERT (Wu *et al.*, 1999).

In order to construct an adenoviral vector for detecting circulating tumor cells, we used the core promoter as reported in a publication from Song (Fig. 22) (Song, 2004).

1	AGTGGATTTC GGGGCACAGA CGCCAGGAC CGCGCTTCCC ACGTGGCGGA GGGACTGGGG ACCCGGGCAC CCGTCCTGCC CCTTCACCTT CCAGCTCCGC
	TCACCTAAGC GCCCGTGTCT GCGGGTCTTG GCGCGAAGGG TGCACCGCCT CCTTGACCCC TGGGCCCGTG GGCAGGACGG GGAAGTGGAA GGTGAGGGCG
101	CTCCTCCGCG CGGACCCCGC CCGTCCCGA CCCCTCCCGG GTCCCGGGCC CAGCCCCCTC CGGGCCCTCC CAGCCCCCTC CCTTCCTTTC CGCGGCCCGG
	GAGGAGGCGC GCTTGGGGCG GGGCAGGGCT GGGGAGGGCC CAGGGGCCGG GTCGGGGGAG GCCCGGGAGG GTCGGGGAGG GGAAGGAAAG GCGCCGGGGC
201	CCCTCTCCTC GCGGCGCGAG TTTCAGGCAG CGCTGCGTGG CGGCCGCAAA ACTGCTGCGC ACGTGGGAAG CCCT
	GGGAGAGGAG CGCCGCGCTC AAAGTCCGTC GCGACGCAGC GCCGGCGTTT TGACGACGCG TGCACCCTTC GGGA

Figure 21: hTERT promoter nucleotide sequence, according to Accession number AF097365

This region includes all important motifs, like E-boxes, needed for adequate telomerase transcription in our constructed vector.

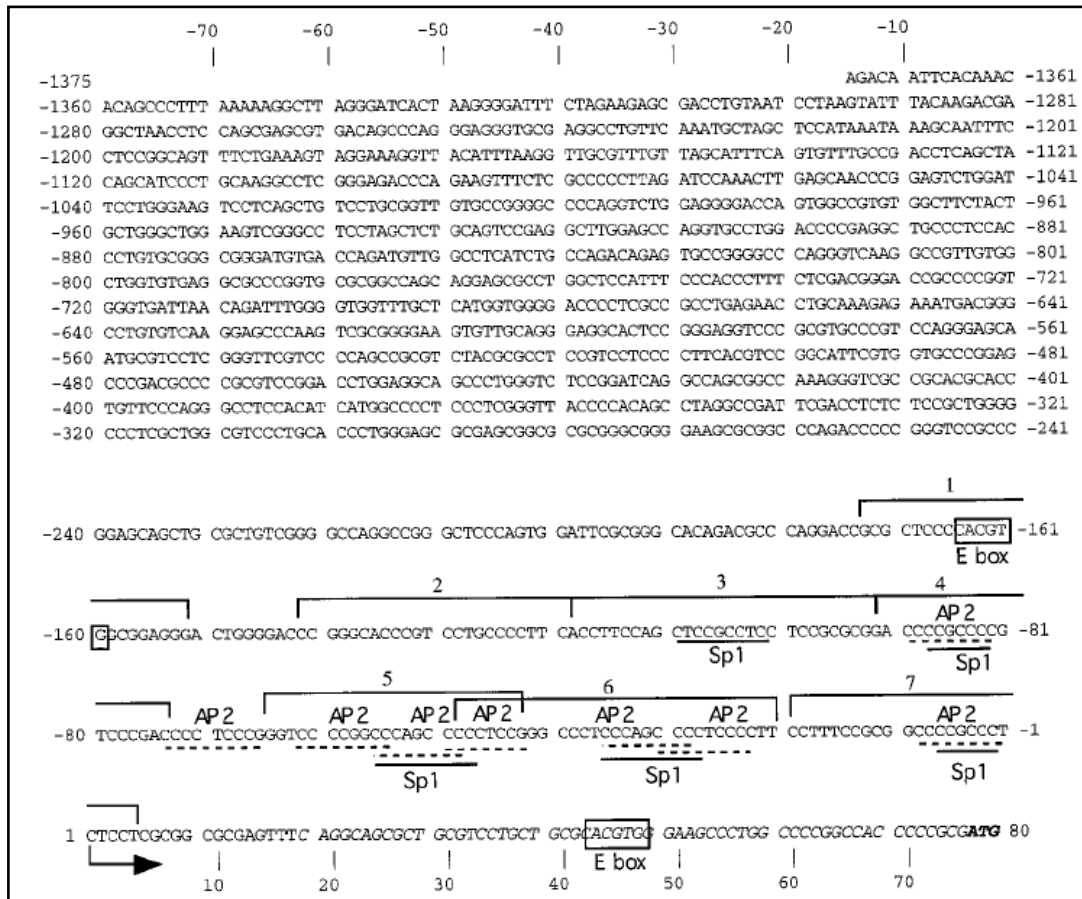


Figure 22: hTERT promoter sequence (Source: Takakura *et al.*, 1999). -1 indicates the transcription start site, whereas +1 displays the initial mRNA-nucleotides. Consensus motifs for E-boxes are indicated by boxes. Consensus motifs for Sp1 and AP2 are underlined.

7.5. Adenovirus

Adenoviruses belong to the family *adenoviridae* and were firstly described in 1953 when W.P. Rowe and his colleagues isolated the virus from adenoid tissue. Today more than 50 different human adenoviral serotypes are known and separated in six subgenera, A – F, according to their molecular and biological properties (Modrow *et al.*, 2003; Kojaoghlanian *et al.*, 2003).

Adenoviral infection in early childhood gives a lifelong immunity to the specific serotype (Kojaoghlanian *et al.*, 2003). Thus antibodies against adenoviral serotypes Ad5 and Ad2 could be detected in more than 50% of the human population, whereas in 90% humoral immunity exists (Chirmule *et al.*, 1999; Russell, 2000).

7.5.1. Adenoviral structure

The adenovirus has a diameter in the range of 80 – 110 nm. It has a non-enveloped icosahedral capsid-structure and external fiber-like projections out from its shell (Fig. 23) (Kojaoghlanian *et al.*, 2003; Modrow *et al.*, 2003).

The main components of the virus are: hexon protein, which is a homotrimeric protein, penton bases, and fibers. Each fiber consists of a trimer protein-complex building a tail, shaft, and the knob (Modrow *et al.*, 2003)

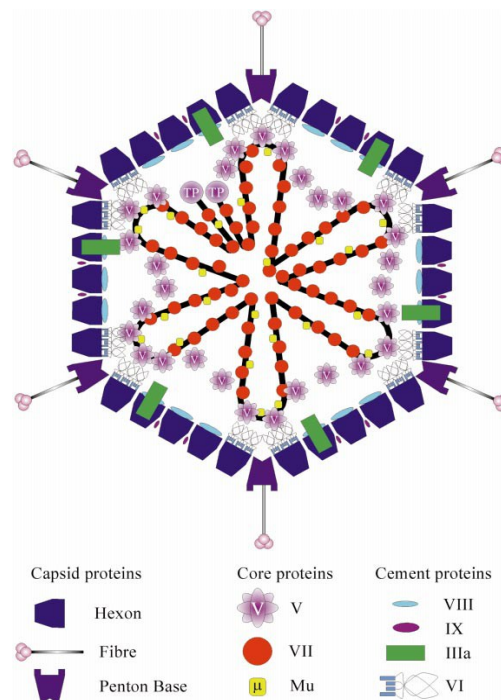


Figure 23: Adenoviral structure (Source: Russel, 2000)

Several “minor” proteins are essential for structure and function of the virus. These proteins (IIIa, VI, VIII, IX) are important in stabilizing the capsid structure and for importing hexon proteins (Vellinga *et al.*, 2005). Other proteins (V, VII, Mu) are located in the virus core and associated with the ds-DNA genome (Russell, 2009).

7.5.2. Adenoviral genome

The adenoviral genome consists of double stranded DNA with a length of 36,000 – 38,000 base pairs. A protein (TP) is bound covalently on both 5' ends of the genome and functions as a primer for replication of adenoviral DNA (Modrow *et al.*, 2003).

The adenoviral genome is separated in different regions according to their temporal point of transcription (Fig. 24) (Tauber & Dobner, 2001):

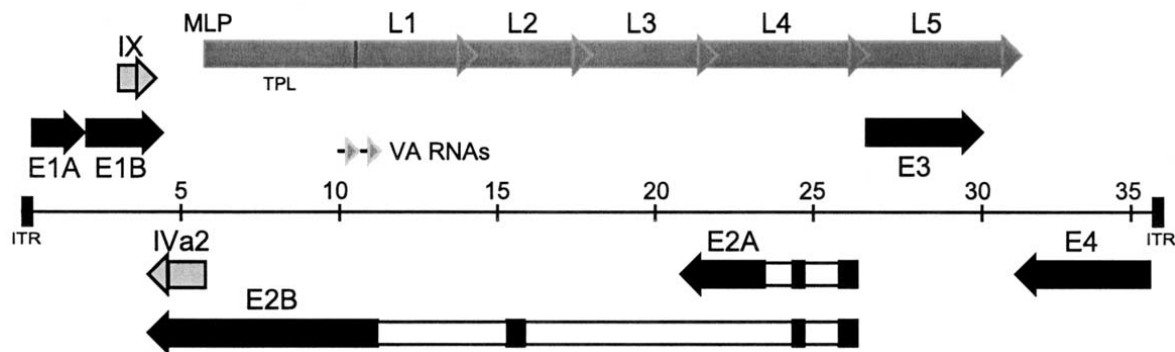


Figure 24: The adenoviral genomic organization (Source: Tauber & Dobner, 2001)

1.) Early transcription region: *E1A*, *E1B*, *E2A*, *E2B*, *E3* and *E4*

These early genes are involved in adenoviral transcription, replication of the viral genome, and host immune suppression.

2.) Delayed transcription units: *IX*, *IVa2*.

3.) Late transcription units: *L1*, *L2*, *L3*, *L4* and *L5*

The late gene products are mostly involved in virus assembly (Breyer *et al.*, 2001; Modrow *et al.*, 2003).

In order to generate a defective adenovirus for gene transfer approaches, the *E1A* and *E1B* genes are replaced by transgenes (Robbins & Ghivizzani, 1998; Breyer *et al.*, 2001). The advantage in removing *E1* is that a transactivation of the viral gene, essential for viral replication, is not feasible (Kay *et al.*, 2001). Since these vectors are characterized by limited packaging capacity and reduced expression of the transgene, these adenoviral vectors have been improved (Breyer *et al.*, 2001).

The second generation of adenoviral vectors lacks *E1* and *E3*; moreover some vectors have been developed with a triple deletion (*E1/E3* and/or *E4*), and these viruses showed a reduced antigenicity (Breyer *et al.*, 2001; Robbins & Ghivizzani, 1998). Since the virus is deficient in *E1* gene, the new generated virus is only able to propagate in a cell-line that provides *E1* gene products, such as HEK 293 cells (Robbins & Ghivizzani, 1998).

Compared to other viral vectors, the use of adenoviruses bears some advantages, since adenoviral vectors are able to infect non-replicating cells (Fujiwara, 2011). However, the

transduction with the viral genome is lost during cell division and - as it is of transient nature - the expression of the “gene of interest” decreases.

7.5.3. Construction of a telomerase-specific vector

The development of cancer-specific modified adenoviruses was a forward breakthrough in the field of cancer research, since the replication of these vectors is restricted to tumor cells.

Three different types of adenoviral-vectors are described in the literature:

- 1.) The deletion of the *E1B* gene in ONYX-015, an adenoviral vector, consequently leads to lysis of p53-deficient tumor cells, whereas “normal” or “healthy” cells will survive (Mathis *et al.*, 2005).
- 2.) In order to restrict adenoviral vector replication to tumor cells, heterologous promoters were used for the regulation of *E1A* gene transcription. Promoters, which were frequently used in this context, are: MUC1, prostate-specific-antigen, or osteocalcin. The specificity of this approach is critical, since the composition of the primary tumor is heterogeneous (Kawashima *et al.*, 2004).
- 3.) The adenoviral vector OBP-401 was designed to ensure tumor-cell-specific replication. The process of replication depends on telomerase-activity of the cell and is controlled by the hTERT-promoter which drives the transcription of *E1A* gene, leading to efficient lyses of tumor cells (Deng *et al.*, 2002). Although it was reported that *E1A* expression might be sufficient to kill tumor cells (Li *et al.*, 2001), *E1A* and *E1B* are connected to each other by IRES to improve the specificity of OBP-401 (Kawashima, 2004).

In order to ensure the monitoring of viral replication and to further visualize viral distribution, the green fluorescent protein (GFP) - driven by the CMV-promoter - was inserted into the viral vector (Fujiwara *et al.*, 2006) and infected tumor cells are detectable due to the GFP expression.

7.5.4. Detection of circulating tumor cells using telomerase-specific adenovirus

Current techniques for the detection of circulating tumor cells (CTCs) were described in chapter 1.8. Since more than 85% of all cancer cells display telomerase activity, expression of this enzyme seems to be a promising marker for the detection of tumor cells (Wirth *et al.*, 2003; Fong *et al.*, 2009). Kawashima *et al.* constructed the viral vector OBP-301 (Telomelysin), in which *E1A* and *E1B* genes are under the control of hTERT promoter (Fig. 25) (Kawashima *et al.*, 2004; Fujiwara, 2011). It has been shown previously that the linkage

of E1A and E1B by IRES (internal ribosome entry site) has the potential to increase the specificity of the virus (Li *et al.*, 2001). Hence, in OBP-301, the adenoviral replication is restricted to telomerase-positive cells and these cells are efficiently killed by the virus (Fujiwara, 2011).

In order to facilitate a tumor-specific imaging Umeoka and colleagues used the green fluorescent protein (GFP), which originates from the jellyfish *Aequorea Victoria*, in combination with the replication-deficient adenovirus OBP-301 (Umeoka *et al.*, 2004). Since the expression of GFP is driven by the cytomegalovirus (CMV) promoter (Fig. 25), the new construction allows monitoring of viral replication. This new created adenovirus, termed OBP-401, labels tumor cells *in vitro* and also *in vivo* by expressing the green fluorescence protein (Kishimoto *et al.*, 2006).

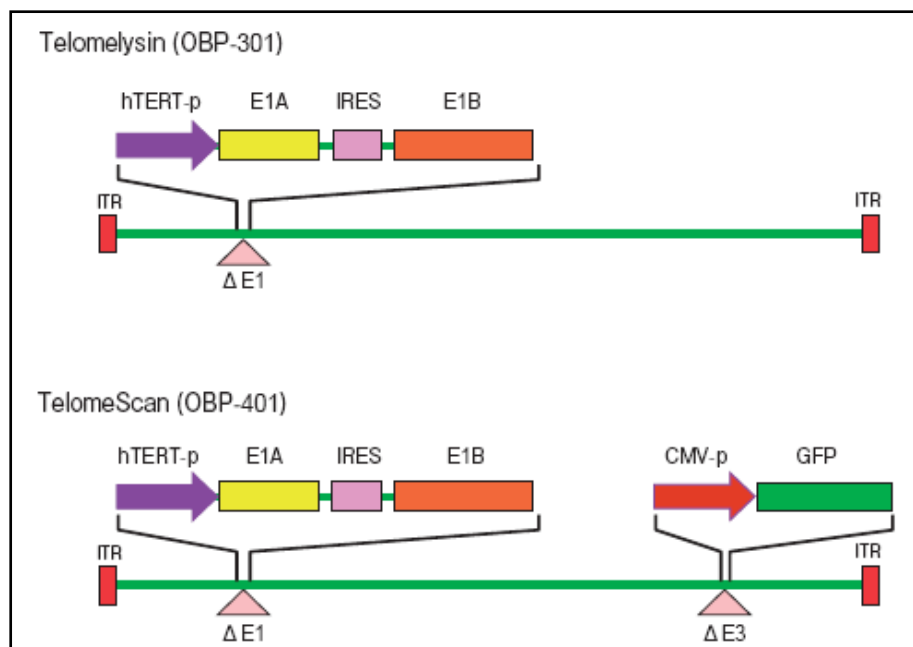


Figure 25: Telomerase-specific oncolytic adenoviruses (Source: Fujiwara, 2011)

In OBP-301, E1A and E1B are linked with an IRES-sequence and hTERT promoter drives their expression. In OBP-401, GFP is inserted in E3 region and expressed under the CMV promoter.

In addition, a new protocol was created to selectively identify fluorescent labeled human CTC among millions of peripheral blood cells. This novel and simple approach involves three steps:

- 1.) Red blood cell lyses
- 2.) Infecting the cell pellet with modified adenovirus
- 3.) Detection of CTC with a fluorescence microscope (Kojima *et al.*, 2009)

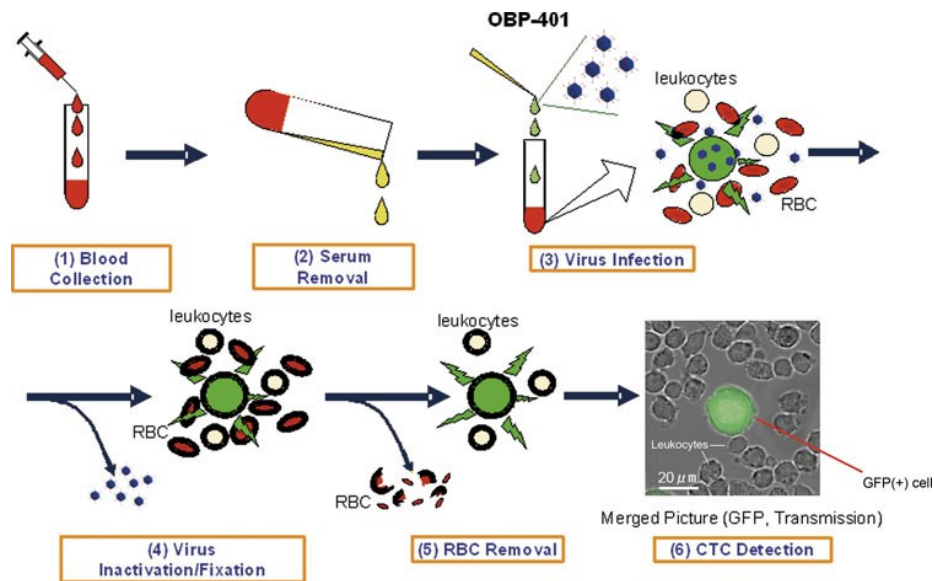


Figure 26: Detection of CTC with OBP-401 assay. Peripheral blood (7,5ml) is drawn from the cancer patient, serum is removed, and the blood is washed with PBS. Blood cells are incubated with OBP-401 for 24h at 37°C. The sample is fixed with paraformaldehyd and GFP-positive cells are detected using fluorescence microscopy (Kim *et al.*, 2011).

It has been demonstrated that OBP-401 is comparable to other methods, such as CellSearch system, regarding their ability to detect CTCs. During the process of vascular invasion of cancer cells, epithelial-mesenchymal-transition (EMT) plays an important role. These cells undergoing EMT show a low expression of surface marker as EpCAM and are not detected by the CellSearch System. This problem could be solved using OBP-401 for CTC detection. Since telomerase is expressed autonomously, (Fig. 26) CTCs can be easily identified by the telomerase-specific virus (Kim *et al.*, 2011).

8. AIMS OF THE PROJECT

8.1. General aims

Construction of an adenoviral telomerase-specific vector:

In order to detect circulating tumor cells in the peripheral blood of cancer patients, the establishment of an adenovirus specific for telomerase positive cells was planned.

Activity of the telomerase enzyme is specific for cells of the immune system, fetal cells, but also cancer cells. In the course of this project we attempted to establish a viral vector construct, harboring the hTERT promoter linked GFP.

The hypothesis of this part of the thesis was that the identification of CTCs, which is currently extremely difficult (see part I of the thesis), can be greatly facilitated by infection of the peripheral blood by the modified adenovirus, because telomerase-transcription factors will bind the promoter region of telomerase-positive cells and switch on the expression of GFP.

8.2. Specific aims

First, we planned to construct a telomerase-specific pShuttle2/hTERT/GFP vector in an adequate bacterial strain.

In a second step we investigated telomerase expression in different cancer cell lines by quantitative Real-Time-PCR.

Additionally, we performed a functionality assay by applying the constructed vector on different cancer cell lines, in order to monitor hTERT expression in living cells.

9. Methods

9.1. Bacteria work

9.1.1. Overnight culture

In order to enrich the bacteria 5 ml of sterile LB media containing antibiotics were inoculated with one single bacterial colony. To ensure that all used materials were sterile a dummy tube was inoculated with no cells in order to verify that the medium, tubes, and sticks were sterile. Infected media was incubated overnight at 37°C and with shaking at 180 rpm. Bacterial culture was used for further plasmid isolation.

9.2. DNA purification methods

9.2.1. GenElute® Plasmid Miniprep Kit (Sigma Aldrich)

2 ml of an overnight *E. coli* culture was centrifuged for 1 min at 12 000 x g to pellet the cells. Supernatant was carefully removed. The bacterial pellet was completely suspended in 200 µl of resuspension buffer. By adding 200 µl of lysis buffer cells were lysed. Immediately after that cells and buffer had to be mixed by inverting the tube 4-6 times. It is important that the lysis reaction does not exceed 5 min because a prolonged alkaline lysis might permanently denature supercoiled plasmid DNA and this would be unsuitable for downstream applications. In a next step the cell debris were precipitated by adding 350 µl of Neutralization/Binding Solution. For mixture of the buffer and cells tubes had to be inverted 4-6 times followed by a centrifugation step 10 min at 12 000 x g. The Binding Column was inserted into a microcentrifuge tube, 500 µl Column preparation solution was added on each and centrifuged for 1 min at 12 000 x g. Flow-through was discarded and 500 µl wash Solution were added on silica-membrane to remove residual salt and other contaminants. The DNA was eluted from the silica-membrane by adding 50 µl nuclease free water (preheated at 55°C) followed by centrifugation for 1 min at 12 000 x g. For plasmids with high copy number concentrations in the range of 50 - 80 ng/µl were obtained.

9.2.2. GenElute[®] PCR Clean Up Kit

For purification of DNA e.g.: after PCR reactions, digestion with restriction enzymes or dephosphorylation, the GenElute[®] - PCR Clean Up Kit (Sigma GenElute Kit, Sigma Aldrich Chemie GmbH, NA1020) was used. The kit is designed for rapid purification of single-stranded or double-stranded PCR amplification products (100 bp to 10 kb) from other components. DNA is bound on a silica membrane and eluted in 50 µl Elute buffer diluted 1:10 in water, preheated at 55°C.

9.2.3. Purification of DNA extracted from agarose gel

The use of Sigma GelExtraction Kit offers rapid purification of DNA (linear or Plasmid-DNA) in the range of 50 bp to 10 kbp from low-melting agarose (e.g.: low gelling temperature, Sigma Aldrich). The kit combines a silica-binding technology with convenience of a spin or vacuum based format.

First the DNA fragments have to be cut from the gel and weighed. The slices of an agarose gel were solubilised with 3 gel volumes of the Gel Solubilization Solution. This solution dissolves gel slices from agarose based gels run in either TBE or TAE buffers. With the pH-indicator a visualization of the gel slices works easily and the colour changes indicate whether the pH of the sample is optimal for binding on the membrane or not. For example, if the fragment has 100 mg, 300 µl of the Gel Solubilization Solution were added. The mixture was incubated at 55-60°C for 10 min and has to be vortexed every 2-3 min in order to dissolve the gel fragment completely.

Meanwhile the columns were prepared to ensure the binding ability of the membrane. Therefore 500 µl Column Preparation Solution were put on the membrane and tubes were centrifuged for 1 min at 12 000 x g. Flow-through was discarded.

Once the gel slices are dissolved the colour of the mixture has to be checked. If the colour of the sample is yellow, sample can be put on the membrane and purification can be proceeded. If the sample is red, adding of 10 µl of the 3 M Sodium Acetate Buffer, pH 5.2, is necessary for binding of the sample on the membrane. Due to the acid character of Sodium Acetate color will switch to yellow and samples can be processed.

One gel volume of isopropyl alcohol was added and the sample was mixed. In the next step, extracted DNA fragments were selectively absorbed onto a silica membrane followed by a centrifugation step for 1 min at 16 000 x g.

The flow-through was discarded and the column washed with 700 µl Wash Solution. After centrifugation of the sample, the flow-through was discarded again and in order to dry the membrane, the filter was placed back into the collection tube and centrifuged for another min at 16 000 x g. Finally, the DNA was eluted in 50 µl distilled water, which previously had been heated at 55°C to increase recovery.

9.3. DNA amplification methods

9.3.1. Polymerase Chain reaction (PCR)

Fibroblastic cells containing a plasmid for telomerase over-expression (BJ hTERT shp53 pBABE) were used as template for hTERT-promoter amplification, whereas a commercially available plasmid pAcGFPN1 (Clontech) acts as template to obtain the GFP-PCR-product. Specific primers were designed to attach specific restriction sites to the fragments.

The PCR reaction was performed according to the Phusion High Fidelity Polymerase protocol. Phusion DNA Polymerase (Finnzymes OY, Finland; F-530S; 100U) have a 5'→3' DNA polymerase activity and 3'→5' exonuclease activity. This generates blunt ends in the amplification products. The polymerase is also capable of amplifying long amplicons such as the 7.5 kb genomic and 20 kb plasmid DNA.

The chosen temperature protocol correlates with the length of the fragment, the primer annealing temperature but also the GC content and length of the primers.

The denaturation step in this protocol is 10 seconds at 98°C, primer annealing was performed at 63°C for 30 seconds and extension time was 30 seconds with 72°C. After 35 cycles of the PCR reaction a final elongation step with 72°C for 10 min finish the reaction. Finally amplified products were cooled to 8°C.

9.3.2. Colony-PCR

In order to test whether the transformed colonies contain the “insert of interest” or if the colonies harbour a relegated plasmid, colony-PCR was performed.

For this method the HotStarTaq PCR Master Mix Kit (Qiagen GmbH, Hilden, Cat.No: 203445), primers in a concentration of 10 µM for the insert of interest and water were mixed to a Master Mix. 50 µl of the mixture were aliquoted to reaction tubes and one single bacterial colony transferred to the tube and were suspended in the mixture. The bacterial colony

functioned as template in this PCR-reaction. The temperature program for PCR reaction was done according to the following protocol.

Tab.3: Temperature program for colony-PCR reaction

Initial activation step	15 min	95 °C
Denaturation	1 min	94°C
Annealing	0,5 min	60°C
Extension	1 min	72°C
Number of cycles	30	
Final extension	10 min	72°C

PCR-products were load on a 1.5% agarose gel to estimate size, concentration of the PCR-product and specificity of the primer/ PCT product.

9.4. Cloning Methods

9.4.1. Restriction Digestion

Restriction enzymes were used to cut the DNA fragments on specific recognition sites, consistent of palindromic sequences (4, 6 or 8 nucleotides).

These enzymes produce either sticky- or blunt ends. For our work, we chose enzymes producing sticky ends in order to ensure best ligation conditions.

For digestion of 1 µg DNA, 35 units of *MfeI*-HF, 200 units *XbaI*, 200 units *KpnI* and 300 units *AflIII* were used.

DNA was digested either for three hours or 16 hours at 37°C. Enzymes were then inactivated for 20 min at 65°C.

9.4.2. Ligation

Production of phosphodiester bridges between the 5'-phosphate of one DNA donor molecule and the 3'-hydroxyl group of a DNA acceptor molecule is catalysed by the enzyme ligase.

For ligation between the purified vector and an insert of interest, both DNA fragments have to be in a ratio of 1:3. Depending on the size ratio from vector to insert, 150 ng - 200 ng of the vector-DNA was ligated with 100 - 180 ng DNA from the insert.

Usually a ligation reaction from 20 µl consists of 1 µl T4-ligase, 2µl of T4-buffer, DNA (from vector and insert) and water. The reaction was incubated for two hours at room temperature or overnight at 16°C.

9.4.3. CloneJet™ PCR Cloning Kit

The CloneJET™ PCR Cloning Kit (Fermentas) is a positive selection system for efficient cloning of PCR products generated by polymerases generating blunt-ends (High Fidelity Polymerase, Phusion) but also 3'-dA overhangs (*Taq*-DNA Polymerase).

Additionally, any other DNA fragment, either blunt or sticky-end, can be successfully cloned using this kit.

pJET1.2/blunt contains a lethal gene which is disrupted by ligation of a DNA insert into the cloning site. As a result, only cells with recombinant plasmids are able to propagate, eliminating the need for expensive blue/white screening.

(Fig. 27)

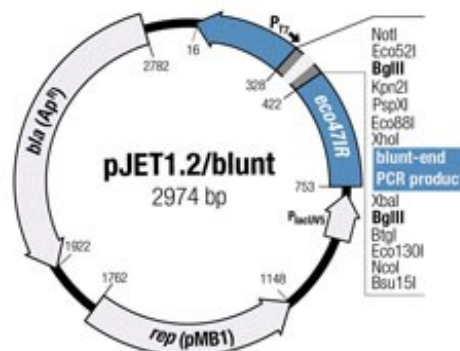


Figure 27: Vector map: pJet/blunt 1.2
(Source: www.thermoscientific.com)

PCR fragments of hTERT promoter and GFP were ligated into the pJET vector for 20 min according to the sticky end cloning protocol and transformed into chemically competent *E. coli* DH5 alpha cells.

9.4.4. Transformation of chemically competent cells

Chemically competent *E. coli* DH5alpha cells were transformed according to the manual *One Shot® MAX Efficiency® DH5α™-T1R Competent Cells* (Invitrogen). 50 µl chemically competent cells were thaw on ice for approximately 10 min. One to five µl of each ligation

reaction were directly pipetted into the competent cells and mixed by tapping the tube gently. After that the vial has to be incubated 30 min on ice. Incubation of the tube at 42°C for 30 seconds shocked the cells. In order to allow DNA to enter the cell membrane and to transform the competent cells they have to be heat shocked at 250 µl SOC medium was added to the freshly transformed cells and mixed at 300 rpm for 45 min at 37°C. After regeneration 50 µl and 150 µl of the cells were spread on LB plates containing Kanamycin (50 µg/ml) and incubated over night at 37°C.

9.4.5. Sequencing of genomic and plasmid-DNA

We performed Sanger sequencing to confirm the accuracy of the vector-sequence after all cloning steps.

For each sequencing reaction 7 µl distilled water, 1 µl Big Dye, 1 µl 5x Big Dye buffer and 0.5 µl PCR- product were mixed. The sequencing was performed separately in forward and reverse direction therefore 0.5 µl forward- or reverse primer, in a concentration of 10 µM, were used.

Following PCR program was used for DNA amplification:

Tab. 4: Sanger Sequencing PCR-Program

Temperature (°C)	Time	} 24 x
96	30 sec	
50	15 sec	
60	4 min	
8	forever	

Sequencing reaction products were purified with SEPHADEX G50 columns.

Therefore, the PCR reaction was diluted 1:1 with distilled water.

250 µl Sephadex were transferred into the PCR- stripes and centrifuged for 1 min at 750 rpm to remove the water from the column. The flow-through was discarded and Sephadex-tubes were centrifuged for another min at 750 rpm. After that, Sephadex columns were transferred

into new PCR-strips and diluted PCR reaction was load onto the columns. In a next step tubes were centrifuged for another two min in order to purify the sequencing reaction.

DNA samples were transferred into a sterile 96-well-PCR-plate and samples were analyzed with the help of 3100 sequencer from Applied Biosystems. Finally sequencing data were analyzed by “Seqscape v2.5”.

9.4.6. PCR for hTERT and GFP

In order to check if the insert was successfully ligated into the vector, primers were chosen that anneal onto pShuttle vector, and additionally onto hTERT and GFP sequences.

DNA is amplified according to the Phusion protocol (Finnzymes) and PCR products were checked by agarose gel-electrophoreses. Therefore, 6 µl DNA is mixed with 3 µl Stop-buffer and load on a 1.5% gel.

For further estimation of the fragment size a 100 bp ladder (Invitrogen) run simultaneously on the gel. The PCR-product is used for subsequent sequencing reaction.

9.5. Experiments with human cell lines

9.5.1. Cultivation of human cell lines

Seven out of the ten cell lines which have been used, were cultured under standard culture conditions. All experiments were performed under the laminar flow to avoid contamination of cultivated cells.

Media was changed every second day, splitting of cells was performed depending on doubling time of the cells. Therefore, adherent cells were washed once with preheated (37°C) 1x PBS, in order to remove the media, which would block the trypsin-reaction.

Cell layer was incubated with trypsin until cells started to detach the dish. Finally, adding of 8 ml of serum-containing media blocked the reaction. Cells can be separated and transferred in flasks or be pelleted by centrifugation (e.g. 1500 x g, 5 min).

9.5.2. RNA isolation

We isolated total RNA from human colorectal, breast and HCC cell lines by using RNeasy® Mini Kit (Qiagen), according to the protocol provided by the manufacturer. Total RNA was eluted in 30 µl sterile water and RNA concentrations were estimated by Nanodrop®. Isolated RNA was stored at -80°C.

9.5.3. Reverse Transcription PCR (RT-PCR)

Total RNA was transcribed in complementary-DNA (cDNA) by Omniscript® Transcription Kit (Qiagen) together with Oligo(dT) Primer (Invitrogen). The transcription reaction was performed with 2 µg RNA in 20 µl. cDNA was stored at -20°C until further use.

9.5.4. Transient Transfection

The introduction of foreign nucleic acids into eukaryotic cells with plasmids is termed “transient transfection”. This method is used to achieve a preliminary (or transiently) expression of a foreign gene in the cell. Various methods - such electroporation, chemical reagents, and viruses - are available for transfection of eukaryotic cells. In this work, the transfection was done by FuGENE® HD reagent (Promega), a non-liposomal agent, which transfects cells with a high efficiency and low toxicity.

In order to test the telomerase-specificity of our construct, it has been introduced either in a telomerase-positive cell line (MCF-7) or telomerase-negative cell line (U2-OS). Additionally, the transfection was performed with negative controls (pShuttle2, pShuttle2-hTERT) and a positive control (pAcGFP-N1).

Transfection of before mentioned cells was done according to the protocol which has been found on FuGENE® HD Protocol Database (<http://www.promega.com/de-de/resources/tools/fugene-hd-protocol-database/>). In brief, one day before transfection adherent cells were plated in 6-well plates at a cell density that ensures that cells are not more than 80% confluent on the day of transfection. One hour before transfection, cells were washed with 1 x PBS and provided with fresh growth-media.

Plasmid-DNA (6.6 µg), OptiMEM® media and transfection reagent were mixed by gently pipetting, followed by an incubation step for 10 – 15 min at room temperature. 150 µl of the FuGENE® HD Transfection Reagent/DNA mixture was transferred to each well. Plates were incubated for further 24 – 48 hours.

Expression of GFP was detected using a fluorescence microscope (CellCelector, ALS).

10. Results

10.1. Construction of a telomerase-specific expression vector

Telomerase activity has been detected in about 80% of cancer cells. Since a unique tumor-cell-marker is not available at the moment, we used this molecular specificity for tumor cell identification. For this purpose we constructed an adenoviral vector containing GFP under the control of the hTERT-promoter. Once this modified adenovirus has infected the cells in blood samples from cancer patients, transcription factors will bind on hTERT-promoter sequence and turn on the expression of GFP. As a consequence, only cells harboring active telomerase will be visible in green.

The construction of a telomerase-specific adenoviral vector was done according to the Adeno-X™ Expression System1 manual from Clontech. All cloning steps were performed with the pShuttle2 vector, which was delivered with the kit. The pShuttle2-vector has a CMV-promoter in front of the multi-cloning-site. Since a CMV-promoter is not usable for our tumor-cell-detection-studies, it was replaced by hTERT-promoter (Fig. 28).

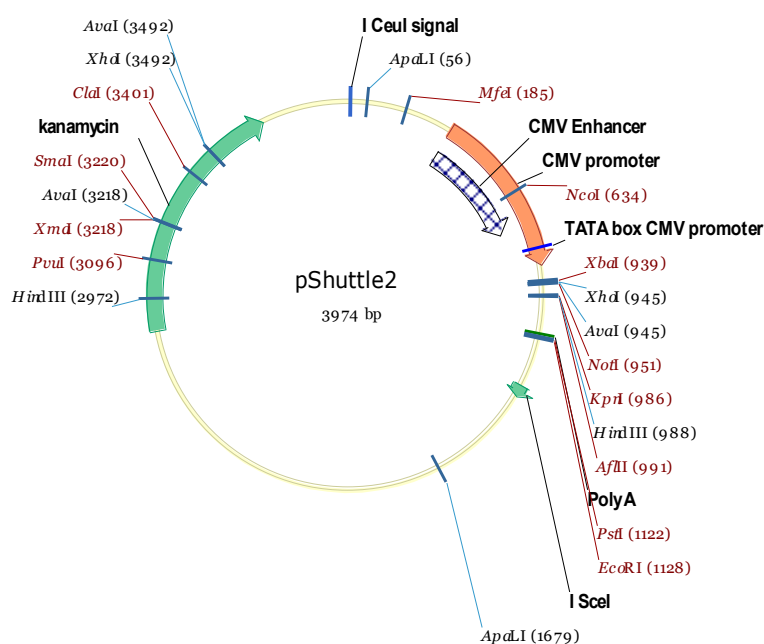


Figure 28: pShuttle2 vector map (clontech). The vector consists of a pUC-ori and harbors a cassette for kanamycin resistance allowing the propagation in *E.coli*. The CMV-promoter and the sequence distal from the MCS are flanked by I-*CeuI* and I-*SceI* recognitions sites, needed for ligation with Adeno-X Viral DNA.

Restriction sites chosen for the promoter replacement had to fulfill the following characteristics:

- They have to be unique within the vector sequence.
- Chosen restrictions sites should not occur within the DNA sequence of the insert.
- The restriction enzyme has to produce sticky-ends (in order to increase the success rate of vector/insert-ligation).

10.1.1. Deletion of the CMV-promoter

For the elimination of the CMV-promoter the pShuttle2-vector was digested with *MfeI* and *XbaI* restriction enzymes. Restriction sites were located on the 5'- and 3' end of the CMV-promoter sequence (Fig. 29).

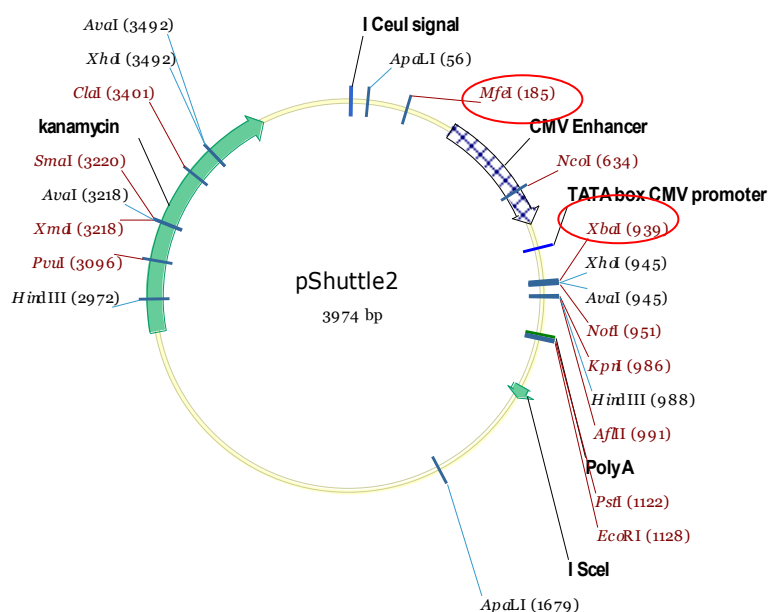


Figure 29: pShuttle2 vector map. (red circles): Restriction sites for CMV-promoter replacement

10.1.2. Amplification of the hTERT-promoter sequence

The hTERT-promoter sequence was amplified with PCR. Genomic DNA from the stable transfected pBABE cell-line was used as a template.

Since cloning of large nucleotide sequences can cause problems, we chose the promoter sequence with the length of 181 bp. (Takakura *et al.*, 1999).

Primers were designed with a nucleotide sequence found in the NCBI- database under the accession number AB016767. These nucleotide sequences were described previously in a study focusing on promoter activity (Song, 2004) and were modified for our purposes by adding restriction sites. For this purpose *MfeI* and *XbaI* restriction sites were fused to the 5' and to the 3' end, respectively, of the hTERT promoter sequence, in order to allow ligation into the pShuttle2 vector.

10.1.3. GFP amplification

The GFP fragment was amplified from the commercially available pAcGFP-N1-vector (Clontech). In the course of amplification, restriction sites for *KpnI* and *AflIII* were attached to the GFP fragment.

After restriction, the digestion GFP fragment was ligated into the new constructed pShuttle2/hTERT-promoter-vector (Fig. 30).

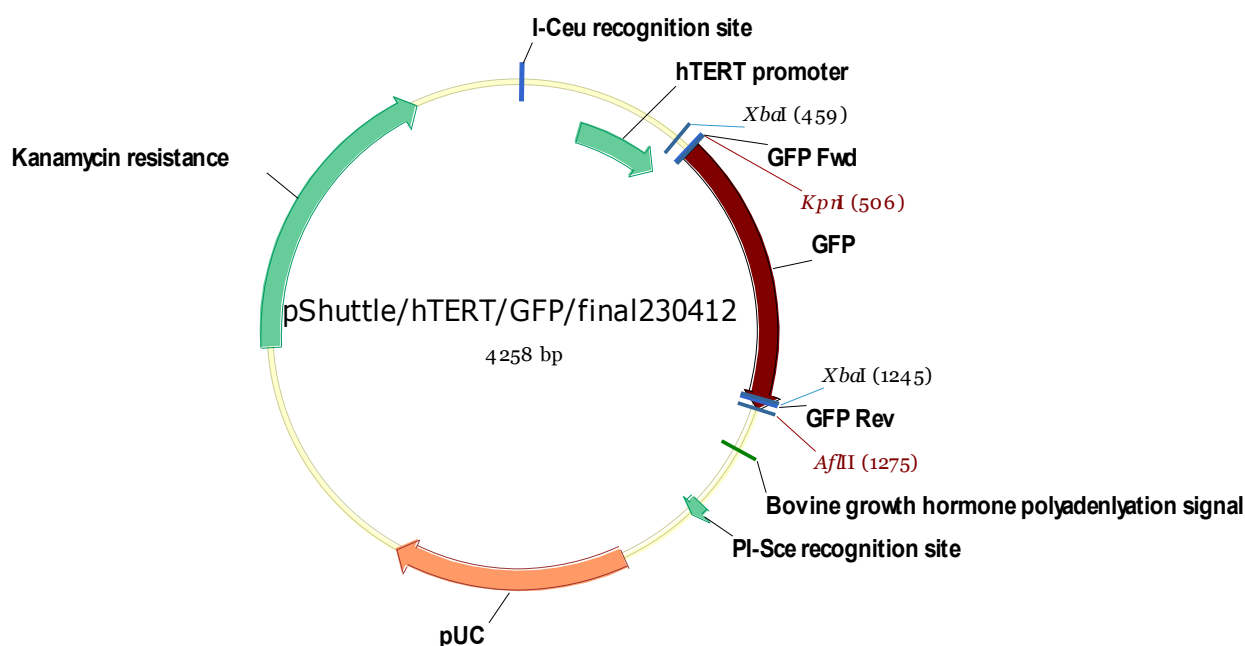


Figure 30: pShuttle2/hTERT/GFP- vector map.

hTERT-promoter was digested with *MfeI* and *XbaI*, and ligated into the pShuttle2-vector. The new constructed vector was propagated in *E. coli*, digested with *KpnI* and *AflIII*. Amplified GFP fragment was also digested with these enzymes and ligated with the pShuttle2/hTERT-promoter.

10.2. Cloning strategy

The telomerase-specific adenovirus was mainly constructed according to the Adeno-X® Expression System 1 protocol from Clontech. Modifications were performed regarding the exclusion of the CMV-promoter in order to replace it by the hTERT-promoter (Fig. 31).

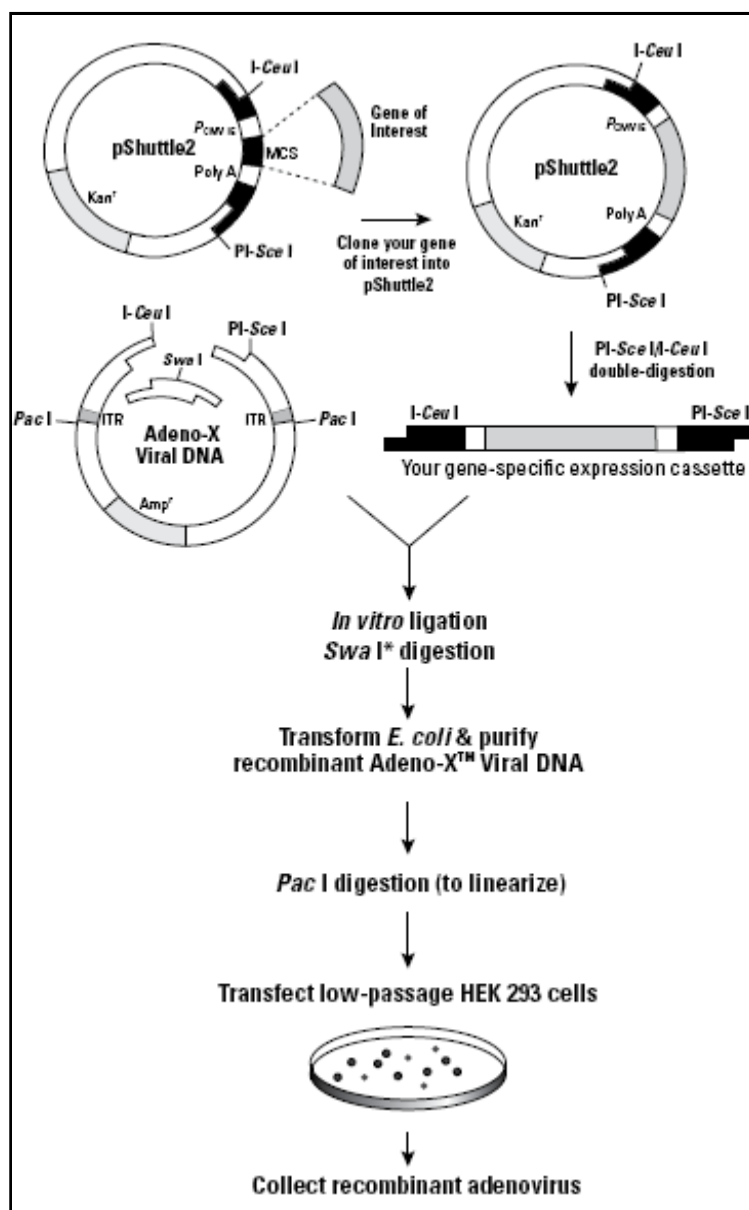


Figure 31: Procedure to construct a recombinant adenovirus using the AdenoX®-Expression system. (Source: Clontech Manual)

10.3. hTERT-Expression

10.3.1. Evaluation of telomerase-activity by qRT-PCR

cDNA from ten different cell lines was used to determine the telomerase-activity by qRT-PCR. 85% of all cancers are characterized by the activity of the enzyme responsible for telomere elongation.

Therefore, telomerase expression was tested in two hepatocellular cancer cell lines (Hep3b, HepG2), two breast cancer cell lines (SKBR-3, MCF7), and three colorectal cancer cell lines (HT-29, HCT116, SW480). Since BJ PD40 cells are fibroblasts and U2-OS cells are tumor cells immortalized by the ALT (alternative lengthening of telomeres)-mechanism, telomerase is consequently not active in these cells. Therefore, BJ PD40 and U2-OS served as a negative control in our qRT-experiments. qRT-PCR was optimized and performed with an annealing temperature of 65°C. Primers were designed according to the nucleotide sequence found on the UCSC Genome browser, in order to amplify fragments in the length of 97 bp and used in a concentration of 250nM. The expression of the housekeeping-gene HPRT (Hypoxanthin Guanin-Phosphoribosyl-Transferase) served as an internal control.

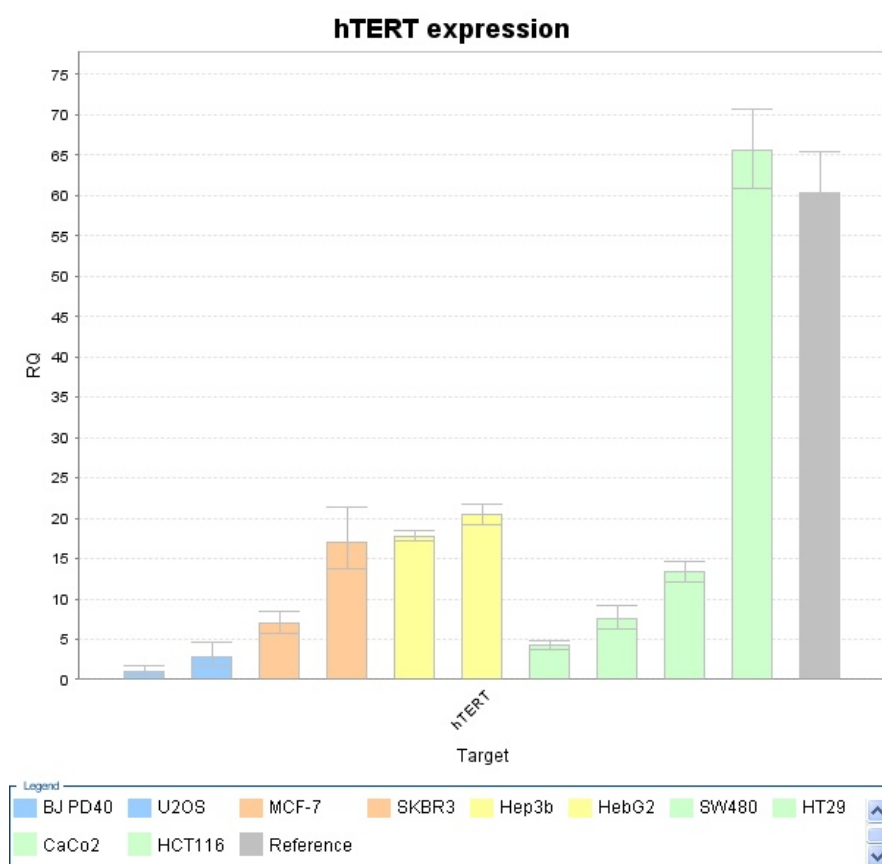


Figure 32: Telomerase expression in different cell lines. All results are shown in relation to BJ PD40 cells, which were set on 1. HCT116 cells showed a significantly higher hTERT expression, compared to all others. Universal Human Reference RNA (pool of 10 cell lines) was used as positive control.

The qRT-PCR results demonstrate that the level of telomerase expression varies between different cell lines, even though telomerase activity has been reported. HCT116 cells showed significantly higher levels of telomerase activity compared to other colorectal cell lines (HT29 or SW 480). In tested hepatocellular cancer cell lines (Hep3b, HepG2) and breast cancer cell lines (SKBR-3, MCF7), telomerase-expression could be detected. The expression levels showed no significant differences within these cell lines.

Furthermore, we could confirm that telomerase is not expressed in fibroblasts (BJPD40) and in osteosarcoma cells (U2-OS). The bars displayed in Fig. 32 are the products of misannealed primers, which were used within this experiment.

10.3.2. Testing the functionality of our new constructed telomerase-specific vector by transient transfection of different cell lines

With respect to the theory that telomerase-positive cells will switch on expression of GFP, we transfected different human cancer cell lines with pShuttle/hTERT/GFP in order to get information about the functionality of the hTERT-promoter.

As confirmed by qRT-PCR, telomerase is not expressed in U2-OS cells, while activity of the enzyme is detectable in MCF7 and HCT116 cell lines. On the other hand, in HCT116 telomerase expression is increased compared to the activity in MCF 7. Since we were interested if changes in the telomerase expression level are reflected in the intensity of GFP expression, we transfected HCT116 (Fig. 33), MCF 7 (Fig. 34) and U2-OS (Fig. 35) with telomerase-specific adenoviral vector.

For this purpose, cells were grown in 6-well-plates to a confluency of 70 – 80%. For transfection FuGene HD® transfection reagent (Promega) and OptiMEM-medium (Invitrogen) were mixed with vector-DNA (6.6 µg, Fugene® HD protocol Database) and incubated with cells of interest for approximately 24 hours at 37°C. In all experiments pAcGFP-N1 vector (Clontech) served as positive control. pACGFP-N1 is characterized by expressing GFP constitutionally under the control of the CMV-promoter.

In addition to our pShuttle2/hTERT/GFP-construct, we deployed other non-GFP-containing-vectors (pShuttle2 and pShuttle2-hTERT) for cell transfection in order to exclude false-positive results.

After 24 hours cells were analyzed with a fluorescent microscope (Olympus CK X41) and pictures were taken (corresponding software) for further analysis.

10.3.2.1. Transfection of telomerase positive HCT 116 cells

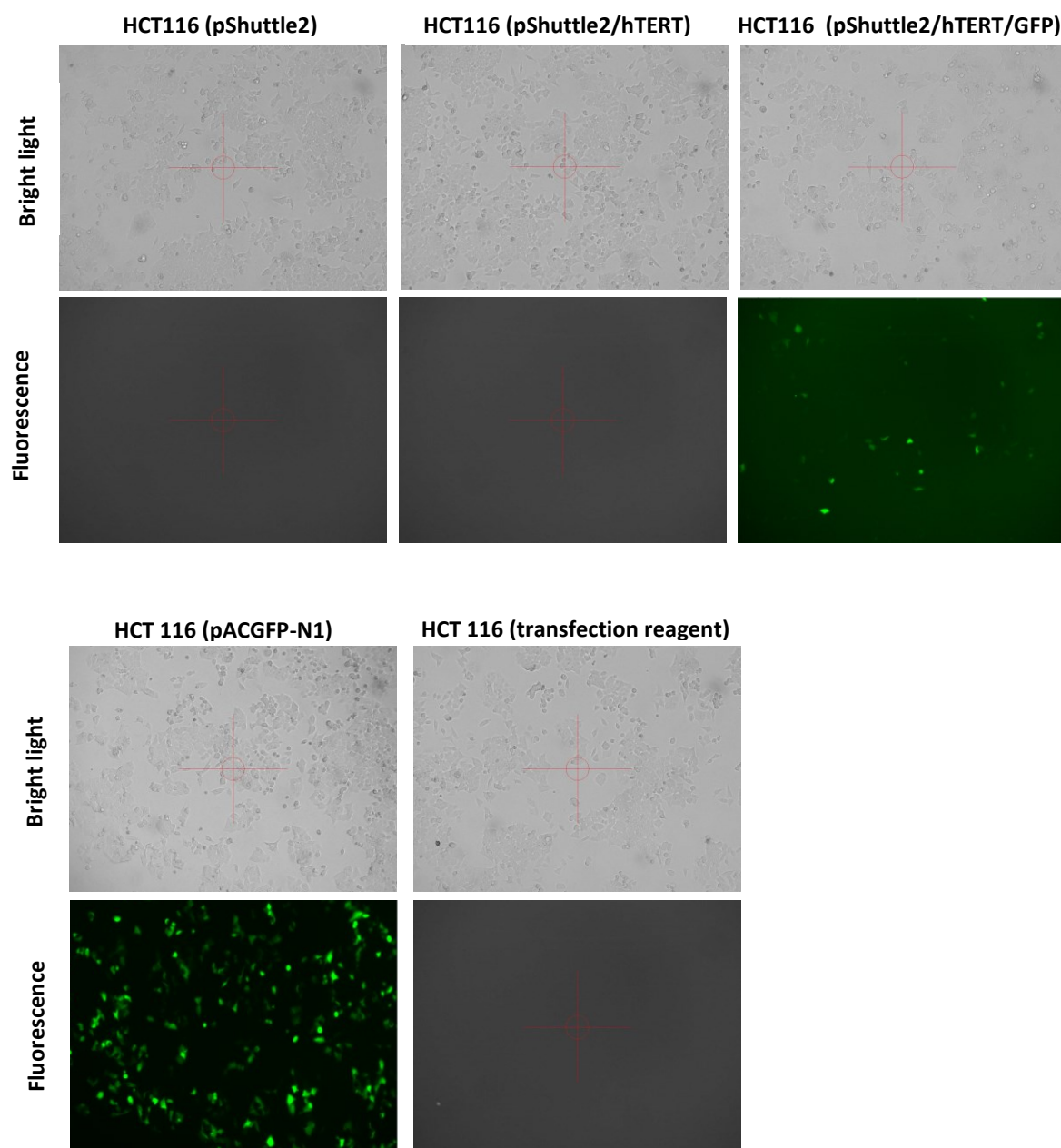


Figure 33: Transfection of HCT116 cell line. Cells were seeded at 80% confluency 12 hours prior to transfection. Transfection was done according to the FuGENE HD® manual for 24 hours at 37 °C. Pictures were made in 10x magnification in bright light and under fluorescence. **pShuttle**, **pShuttle/hTERT** and **transfection reagent** were used as negative controls. **pShuttle2/hTERT/GFP** was transfected in order to test the specificity of our modified vector. **pAcGFP-N1** acts as positive control in this experiment, since telomerase is constantly expressed under the CMV-promoter.

10.3.2.2. Transfection of telomerase-positive MCF7 cells

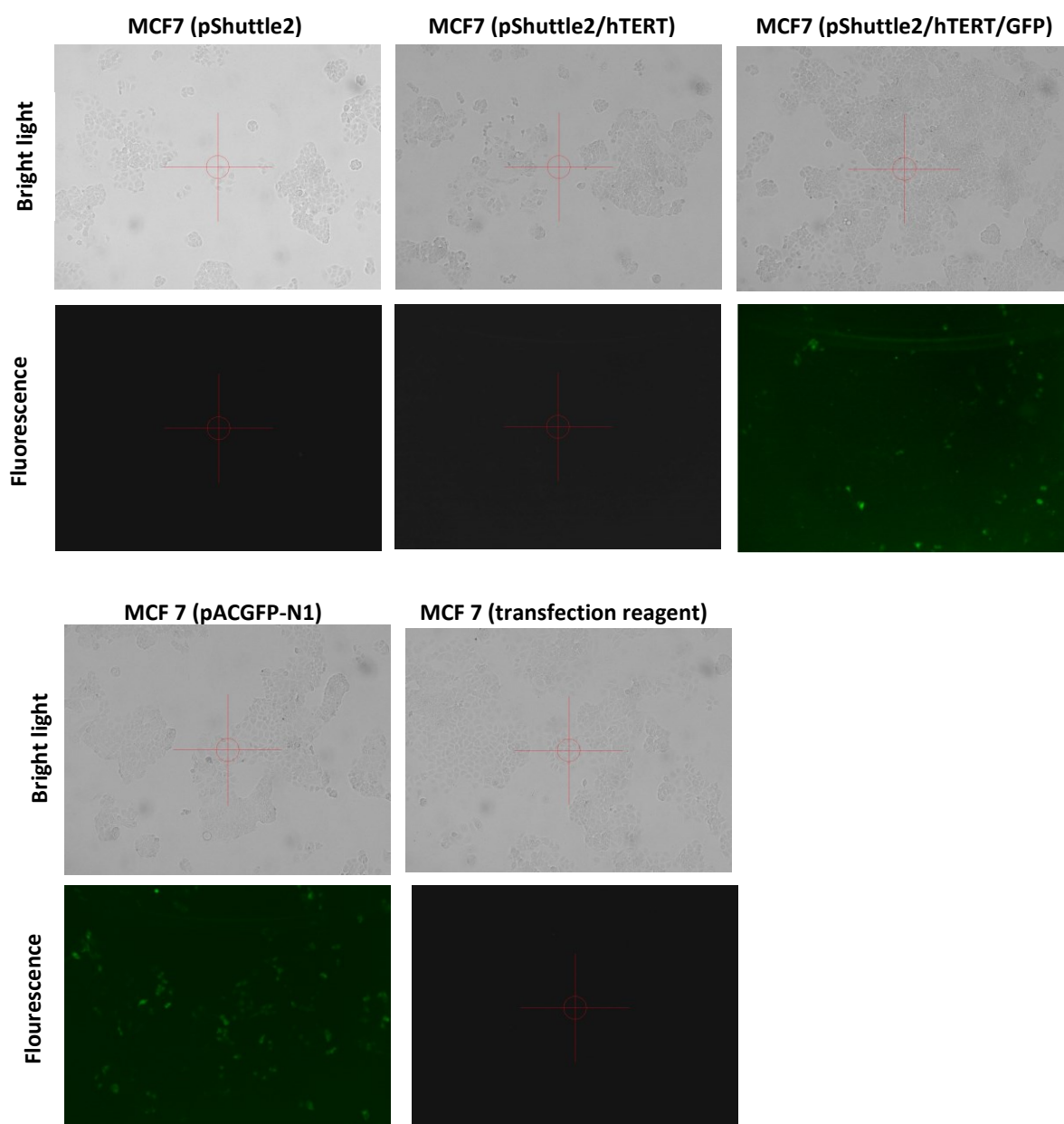


Figure 34: Transfection of MCF7 cell line. Cells were seeded at 80% confluency 12 hours prior to transfection. Transfection was done according to the FuGENE HD® manual for 24 hours at 37 °C. Pictures were made in 10x magnification in bright light and under fluorescence. **pShuttle**, **pShuttle/hTERT** and **transfection reagent** were used as negative controls. **pShuttle2/hTERT/GFP** was transfected in order to test the specificity of our construct. **pAcGFP-N1** acts as positive control.

10.3.2.3. Transfection of telomerase-negative U2-OS cells

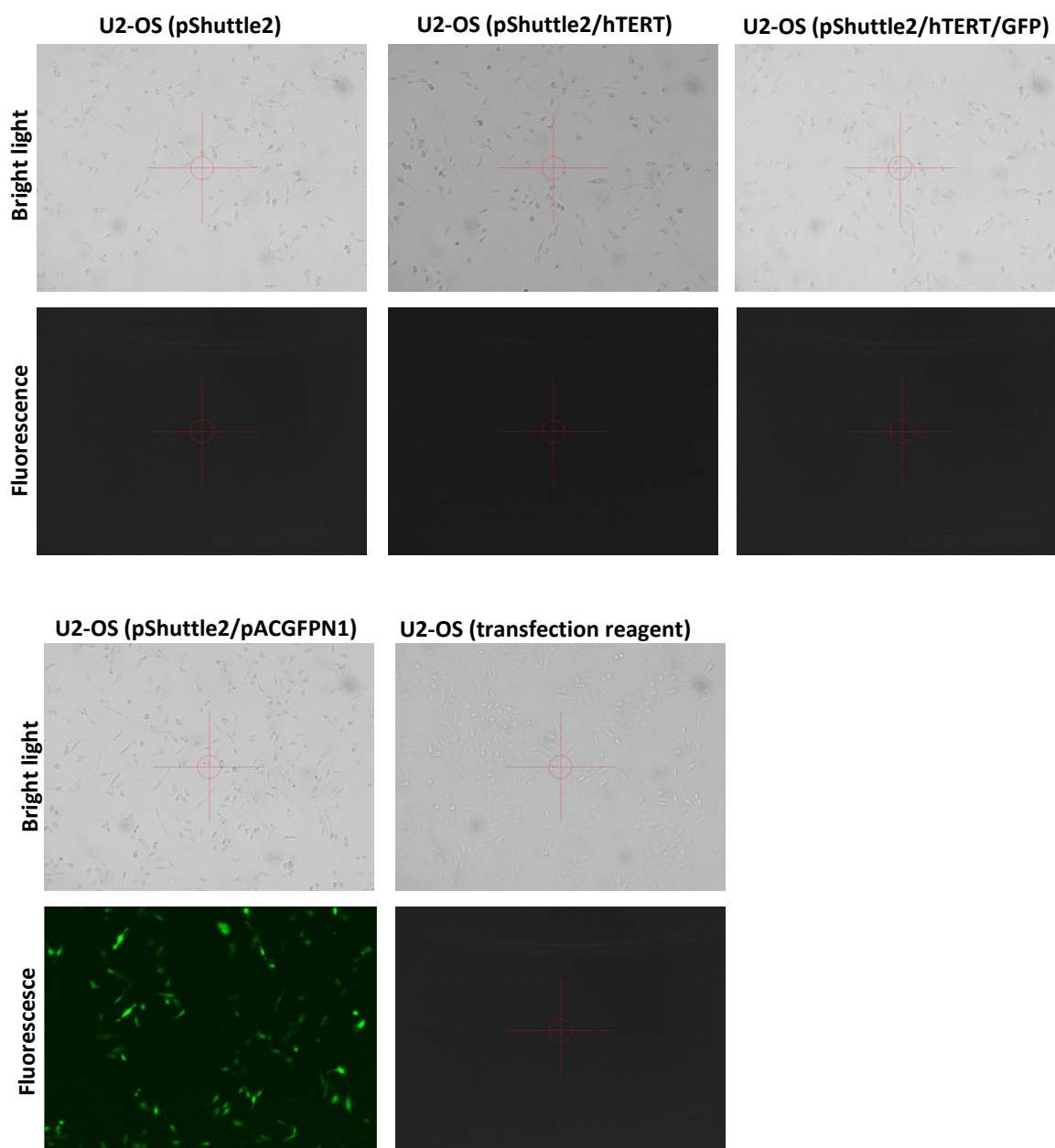


Figure 35: Transfection of U2-OS cell line. Cells were seeded at 80% confluency 12 hours prior to transfection. Transfection was done according to the FuGENE HD® manual for 24 hours at 37 °C. Pictures were made in 10x magnification in bright light and under fluorescence. **pShuttle**, **pShuttle/hTERT** and **transfection reagent** were used as negative controls. **pShuttle2/hTERT/GFP** was transfected in order to test the specificity of our construct. **pAcGFP-N1** acts as positive control.

11. Discussion

This project focuses on the construction of telomerase-specific adenoviral vectors with the capability of detecting telomerase expressing cells, as found in circulating tumor cells. Current systems for CTC detection rely either on cytokeratin characteristics (CellSearch/Veridex-System; CTC-chip) or on the hypothesis that circulating tumor cells have an increased size or different shape (ISET) compared to other blood cells. To date, no specific approaches to specifically detect viable CTCs are available.

A new method, which may in fact facilitate the detection of human CTCs among millions of peripheral blood leukocytes, was described recently. This approach employs a telomerase-specific replication selective adenovirus expressing GFP (Kojima *et al.*, 2010). The system involves a modified GFP expressing adenovirus in which the telomerase promoter (hTERT) has regulated viral replication. Consequently, infection with the telomerase specific virus increases the signal-to-noise ratio, because the fluorescent signal is amplified by viral replication only in viable, infected human telomerase expressing cells (Kojima *et al.*, 2010).

In the course of the project we have established our own vector, which is based on the concept published by Kojima *et al.* (2010), but has some significant modifications in terms of virus construction. Since this version of an adenoviral vector causes specific “oncolysis” - meaning tumor-cells will die after viral infection - OBP-401 is not applicable for our single-cell-analysis approaches. Therefore, we constructed a genetically modified non-lytic telomerase-specific viral-vector, in accordance with the study from Yin and colleagues (Yin *et al.*, 2004), where they described the development of a lentiviral-vector-system (pLNT-EGFP) harboring an expression cassette consisting of hTERT-promoter and GFP for the detection of telomerase-positive tumor cells in cell culture experiments. We modified this vector-system using an adenovirus serotype 5 instead of the lenti-virus. In this context the application of an adenoviral vector bears some advantages in terms of cell infection. In comparison to other types of vectors, e.g.: retroviral vectors, adenoviral vectors are able to infect quiescent and dividing cells. Additionally, they do not integrate into the host genome. Furthermore, adenoviral vectors are generated easily and can be produced at very high titers.

Another modification of the constructed viral vector was the replacement of the CMV-promoter in the backbone of the commercially available pShuttle2-vector by the 181bp nucleotide long hTERT-promoter. Kojama and colleagues succeeded in labelling telomerase active cells (circulating tumor cells) by a telomerase specific adenoviral vector. Hence, in

contrast to the OBP-401 approach, the aim of our virus was not to kill infected cells, and therefore we neglect the insertion of the E1A-IRES-E1B nucleotide sequence in our construct.

11.1. Evaluation of Telomerase-activity by qRT-PCR

Since telomerase expression has been mainly observed in tumor cells, activity of this enzyme has become a hallmark of cancer.

As we planned to test the specificity of our modified construct by transfecting different cell lines, we had to confirm that telomerase is still active in these cells. Although several methods for telomerase detection are available (e.g. TRAP-assay, Kim *et al.*, 1994), we decided to use a PCR-based approach for our experiment (Ulaner *et al.*, 1998).

Ten tumorigenic cell lines were tested by qRT-PCR and, interestingly, the expression levels differed tremendously. MCF 7, a breast cancer cell line which has often been used as a positive control in telomerase assays (Song, 2004), shows a very low telomerase expression. In contrast, hepatocellular cancer cell lines (Hep3b, HepG2) and the colon carcinoma cell line HCT116 displayed a very high expression level. With this knowledge it was interesting to observe the transcriptional activity of the hTERT promoter in HCT116, MCF 7, and U2-OS cells (as negative control) by transfecting them with our modified adenoviral construct (pShuttle/hTERT/GFP). The results confirmed that pShuttle/hTERT/GFP is applicable for the detection of telomerase expressing cells, since telomerase transcription factors bound onto the hTERT promoter and, as a consequence, the GFP gene was transcribed and expressed.

The transcriptional activity driven by the hTERT-promoter could not be detected in telomerase negative U2-OS cells, as these cells are immortalized via the ALT-mechanism (Lafferty-Whyte *et al.*, 2009).

Since we succeeded in confirming that the modified construct is usable for our approaches, the next steps will be the insertion of the adenoviral vector into the adenoviral genome. These steps will be performed by Verena Rupp, M.Sc.

Advantages of a telomerase specific adenoviral vector for the detection of circulating tumor cells

One advantage of this approach compared to the CellSearch system is that one can detect possible circulating tumor cells without any knowledge regarding the morphology of tumorigenic cells. Working with the CellSearch system, employees have to be trained in order to distinguish “normal” cells from tumor cells. This morphological differentiation often poses

a problem for untrained employees in hospitals and other institutions (Maida *et al.*, 2009). Detecting cells by telomerase-specific adenoviruses could overcome these problems.

Another advantage of detecting cells by their telomerase activity is the fact that telomerase-specific adenoviruses are able to detect cells even if they undergo EMT (epithelial-mesenchymal-transition). During the process of EMT, cancer cells lose their epithelial marker, but their telomerase expression is still activated. Hence, circulating tumor cells which would not be detected by the CellSearch system would express GFP and would therefore be identified by the telomerase specific virus (Kojima *et al.*, 2010).

Furthermore, the identification of tumor cells by the virus would raise a new prospect for tumor dynamics measurement. Monitoring of tumor cells in the course of disease progression could be another conceivable application of the virus (Scher *et al.*, 2009). However, there are also some disadvantages of the virus approach.

On the one hand, the adenovirus we will use for construction of our telomerase-specific-approach belongs to serotype 5, which causes a common cold during the childhood. Hence, most people possess antibodies against the virus (Kojima *et al.*, 2010; Kojaghlanian *et al.*, 2003). In order to avoid an immune reaction against the virus in the course of the experiment, Kim *et al.* developed a time consuming washing procedure (Kim *et al.*, 2011).

On the other hand, telomerase activity in “healthy” cells could become problematic in detecting tumor cells by an adenoviral vector. Maida *et al.* described problems regarding the identification of tumor cells in patients with cervical, endometrial, and ovarian cancer, because normal endometrial epithelial cells harbor a very high telomerase activity (Maida *et al.*, 2009).

For these reasons a combination of different approaches would be the best possibility to detect circulating tumor cells

12 Literature

- (1) Anker P, Mulcahy H, Chen XQ, Stroun M. Detection of circulating tumour DNA in the blood (plasma/serum) of cancer patients. *Cancer Metastasis Rev* 1999;18(1):65-73.
- (2) Atamaniuk J, Ruzicka K, Stuhlmeier KM, Karimi A, Eigner M, Mueller MM. Cell-free plasma DNA: a marker for apoptosis during hemodialysis. *Clin Chem* 2006 Mar;52(3):523-526.
- (3) Begum S, Brait M, Dasgupta S, Ostrow KL, Zahurak M, Carvalho AL, et al. An epigenetic marker panel for detection of lung cancer using cell-free serum DNA. *Clin Cancer Res* 2011 Jul 1;17(13):4494-4503.
- (4) Bekaert S, Derradji H, Baatout S. Telomere biology in mammalian germ cells and during development. *Dev Biol* 2004 Oct 1;274(1):15-30.
- (5) Bignell GR, Greenman CD, Davies H, Butler AP, Edkins S, Andrews JM, et al. Signatures of mutation and selection in the cancer genome. *Nature* 2010 Feb 18;463(7283):893-898.
- (6) Blackburn EH. Structure and function of telomeres. *Nature* 1991 Apr 18;350(6319):569-573.
- (7) Blasco MA, Hahn WC. Evolving views of telomerase and cancer. *Trends Cell Biol* 2003 Jun;13(6):289-294.
- (8) Breyer B, Jiang W, Cheng H, Zhou L, Paul R, Feng T, et al. Adenoviral vector-mediated gene transfer for human gene therapy. *Curr Gene Ther* 2001 Jul;1(2):149-162.
- (9) Campisi J, Kim SH, Lim CS, Rubio M. Cellular senescence, cancer and aging: the telomere connection. *Exp Gerontol* 2001 Nov;36(10):1619-1637.
- (10) Center MM, Jemal A, Ward E. International trends in colorectal cancer incidence rates. *Cancer Epidemiol Biomarkers Prev* 2009 Jun;18(6):1688-1694.
- (11) Chaffer CL, Weinberg RA. A perspective on cancer cell metastasis. *Science* 2011 Mar 25;331(6024):1559-1564.
- (12) Chambers AF, Groom AC, MacDonald IC. Dissemination and growth of cancer cells in metastatic sites. *Nat Rev Cancer* 2002 Aug;2(8):563-572.
- (13) Chan KC, Jiang P, Zheng YW, Liao GJ, Sun H, Wong J, et al. Cancer genome scanning in plasma: detection of tumor-associated copy number aberrations, single-nucleotide variants, and tumoral heterogeneity by massively parallel sequencing. *Clin Chem* 2013 Jan;59(1):211-224.
- (14) Chan SR, Blackburn EH. Telomeres and telomerase. *Philos Trans R Soc Lond B Biol Sci* 2004 Jan 29;359(1441):109-121.
- (15) Chirmule N, Probert K, Magosin S, Qian Y, Qian R, Wilson J. Immune responses to adenovirus and adeno-associated virus in humans. *Gene Ther* 1999 Sep;6(9):1574-1583.
- (16) Choi JJ, Reich CF, 3rd, Pisetsky DS. The role of macrophages in the in vitro generation of extracellular DNA from apoptotic and necrotic cells. *Immunology* 2005 May;115(1):55-62.
- (17) Cristofanilli M, Budd GT, Ellis MJ, Stopeck A, Matera J, Miller MC, et al. Circulating tumor cells, disease progression, and survival in metastatic breast cancer. *N Engl J Med* 2004 Aug 19;351(8):781-791.
- (18) Cruz I, Ciudad J, Cruz JJ, Ramos M, Gomez-Alonso A, Adansa JC, et al. Evaluation of multiparameter flow cytometry for the detection of breast cancer tumor cells in blood samples. *Am J Clin Pathol* 2005 Jan;123(1):66-74.
- (19) Deng J, Kloosterboer F, Xia W, Hung MC. The NH(2)-terminal and conserved region 2 domains of adenovirus E1A mediate two distinct mechanisms of tumor suppression. *Cancer Res* 2002 Jan 15;62(2):346-350.

- (20) Dennis Lo YM. PCR-Based Noninvasive Prenatal Diagnosis Using Fetal Cells in Maternal Circulation. *Methods Mol Med* 1998;16:265-274.
- (21) Dhallan R, Au WC, Mattagajasingh S, Emche S, Bayliss P, Damewood M, et al. Methods to increase the percentage of free fetal DNA recovered from the maternal circulation. *JAMA* 2004 Mar 3;291(9):1114-1119.
- (22) Diaz LA,Jr, Williams RT, Wu J, Kinde I, Hecht JR, Berlin J, et al. The molecular evolution of acquired resistance to targeted EGFR blockade in colorectal cancers. *Nature* 2012 Jun 28;486(7404):537-540.
- (23) Diehl F, Li M, Dressman D, He Y, Shen D, Szabo S, et al. Detection and quantification of mutations in the plasma of patients with colorectal tumors. *Proc Natl Acad Sci U S A* 2005 Nov 8;102(45):16368-16373.
- (24) Diehl F, Li M, He Y, Kinzler KW, Vogelstein B, Dressman D. BEAMing: single-molecule PCR on microparticles in water-in-oil emulsions. *Nat Methods* 2006 Jul;3(7):551-559.
- (25) Diehl F, Schmidt K, Choti MA, Romans K, Goodman S, Li M, et al. Circulating mutant DNA to assess tumor dynamics. *Nat Med* 2008 Sep;14(9):985-990.
- (26) Diehl F, Schmidt K, Durkee KH, Moore KJ, Goodman SN, Shuber AP, et al. Analysis of mutations in DNA isolated from plasma and stool of colorectal cancer patients. *Gastroenterology* 2008 Aug;135(2):489-498.
- (27) Donate LE, Blasco MA. Telomeres in cancer and ageing. *Philos Trans R Soc Lond B Biol Sci* 2011 Jan 12;366(1561):76-84.
- (28) Dotan E, Cohen SJ, Alpaugh KR, Meropol NJ. Circulating tumor cells: evolving evidence and future challenges. *Oncologist* 2009 Nov;14(11):1070-1082.
- (29) Fatouros IG, Destouni A, Margonis K, Jamurtas AZ, Vrettou C, Kouretas D, et al. Cell-free plasma DNA as a novel marker of aseptic inflammation severity related to exercise overtraining. *Clin Chem* 2006 Sep;52(9):1820-1824.
- (30) Fehm T, Banys M. Circulating free DNA: a new surrogate marker for minimal residual disease? *Breast Cancer Res Treat* 2011 Nov;130(1):119-122.
- (31) Fiegler H, Geigl JB, Langer S, Rigler D, Porter K, Unger K, et al. High resolution array-CGH analysis of single cells. *Nucleic Acids Res* 2007;35(3):e15.
- (32) Fong SM, Lee MK, Adusumilli PS, Kelly KJ. Fluorescence-expressing viruses allow rapid identification and separation of rare tumor cells in spiked samples of human whole blood. *Surgery* 2009 Sep;146(3):498-505.
- (33) Forshew T, Murtaza M, Parkinson C, Gale D, Tsui DW, Kaper F, et al. Noninvasive identification and monitoring of cancer mutations by targeted deep sequencing of plasma DNA. *Sci Transl Med* 2012 May 30;4(136):136ra68.
- (34) Fujimoto R, Kamata N, Taki M, Yokoyama K, Tomonari M, Nagayama M, et al. Gene expression of telomerase related proteins in human normal oral and ectocervical epithelial cells. *Oral Oncol* 2003 Jul;39(5):445-452.
- (35) Fujiwara T. A novel molecular therapy using bioengineered adenovirus for human gastrointestinal cancer. *Acta Med Okayama* 2011 Jun;65(3):151-162.
- (36) Geiger TR, Peeper DS. Metastasis mechanisms. *Biochim Biophys Acta* 2009 Dec;1796(2):293-308.
- (37) Geigl JB, Obenauf AC, Schwarzbraun T, Speicher MR. Defining 'chromosomal instability'. *Trends Genet* 2008 Feb;24(2):64-69.
- (38) Geigl JB, Obenauf AC, Waldispuehl-Geigl J, Hoffmann EM, Auer M, Hormann M, et al. Identification of small gains and losses in single cells after whole genome amplification on tiling oligo arrays. *Nucleic Acids Res* 2009 Aug;37(15):e105.

- (39) Giulietti A, Overbergh L, Valckx D, Decallonne B, Bouillon R, Mathieu C. An overview of real-time quantitative PCR: applications to quantify cytokine gene expression. *Methods* 2001 Dec;25(4):386-401.
- (40) Goessl C, Heicappell R, Munker R, Anker P, Stroun M, Krause H, et al. Microsatellite analysis of plasma DNA from patients with clear cell renal carcinoma. *Cancer Res* 1998 Oct 15;58(20):4728-4732.
- (41) Gormally E, Caboux E, Vineis P, Hainaut P. Circulating free DNA in plasma or serum as biomarker of carcinogenesis: practical aspects and biological significance. *Mutat Res* 2007 May-Jun;635(2-3):105-117.
- (42) Greider CW. Telomeres and senescence: the history, the experiment, the future. *Curr Biol* 1998 Feb 26;8(5):R178-81.
- (43) Greider CW. Telomere length regulation. *Annu Rev Biochem* 1996;65:337-365.
- (44) Griffith JD, Comeau L, Rosenfield S, Stansel RM, Bianchi A, Moss H, et al. Mammalian telomeres end in a large duplex loop. *Cell* 1999 May 14;97(4):503-514.
- (45) Gupta GP, Massague J. Cancer metastasis: building a framework. *Cell* 2006 Nov 17;127(4):679-695.
- (46) Hahn WC. Role of telomeres and telomerase in the pathogenesis of human cancer. *J Clin Oncol* 2003 May 15;21(10):2034-2043.
- (47) Harley CB, Futcher AB, Greider CW. Telomeres shorten during ageing of human fibroblasts. *Nature* 1990 May 31;345(6274):458-460.
- (48) Hayflick L. The illusion of cell immortality. *Br J Cancer* 2000 Oct;83(7):841-846.
- (49) HAYFLICK L, MOORHEAD PS. The serial cultivation of human diploid cell strains. *Exp Cell Res* 1961 Dec;25:585-621.
- (50) Heitzer E, Auer M, Hoffmann EM, Pichler M, Gasch C, Ulz P, et al. Establishment of tumor-specific copy number alterations from plasma DNA of patients with cancer. *Int J Cancer* 2013 Jul 15;133(2):346-356.
- (51) Hibi K, Robinson CR, Booker S, Wu L, Hamilton SR, Sidransky D, et al. Molecular detection of genetic alterations in the serum of colorectal cancer patients. *Cancer Res* 1998 Apr 1;58(7):1405-1407.
- (52) Horikawa I, Cable PL, Afshari C, Barrett JC. Cloning and characterization of the promoter region of human telomerase reverse transcriptase gene. *Cancer Res* 1999 Feb 15;59(4):826-830.
- (53) Horsman KM, Hickey JA, Cotton RW, Landers JP, Maddox LO. Development of a human-specific real-time PCR assay for the simultaneous quantitation of total genomic and male DNA. *J Forensic Sci* 2006 Jul;51(4):758-765.
- (54) Jahr S, Hentze H, Englisch S, Hardt D, Fackelmayer FO, Hesch RD, et al. DNA fragments in the blood plasma of cancer patients: quantitations and evidence for their origin from apoptotic and necrotic cells. *Cancer Res* 2001 Feb 15;61(4):1659-1665.
- (55) Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. *CA Cancer J Clin* 2011 Mar-Apr;61(2):69-90.
- (56) Jung K, Fleischhacker M, Rabien A. Cell-free DNA in the blood as a solid tumor biomarker--a critical appraisal of the literature. *Clin Chim Acta* 2010 Nov 11;411(21-22):1611-1624.
- (57) Kamangar F, Dores GM, Anderson WF. Patterns of cancer incidence, mortality, and prevalence across five continents: defining priorities to reduce cancer disparities in different geographic regions of the world. *J Clin Oncol* 2006 May 10;24(14):2137-2150.
- (58) Kan Z, Jaiswal BS, Stinson J, Janakiraman V, Bhatt D, Stern HM, et al. Diverse somatic mutation patterns and pathway alterations in human cancers. *Nature* 2010 Aug 12;466(7308):869-873.

- (59) Kawashima T, Kagawa S, Kobayashi N, Shirakiya Y, Umeoka T, Teraishi F, et al. Telomerase-specific replication-selective virotherapy for human cancer. *Clin Cancer Res* 2004 Jan 1;10(1 Pt 1):285-292.
- (60) Kay MA, Glorioso JC, Naldini L. Viral vectors for gene therapy: the art of turning infectious agents into vehicles of therapeutics. *Nat Med* 2001 Jan;7(1):33-40.
- (61) Kim S, Kaminker P., Campisi, Judith. Telomeres, aging and cancer: In search of a happy ending. *Oncogene* 2002;21:503-511.
- (62) Kishimoto H, Kojima T, Watanabe Y, Kagawa S, Fujiwara T, Uno F, et al. In vivo imaging of lymph node metastasis with telomerase-specific replication-selective adenovirus. *Nat Med* 2006 Oct;12(10):1213-1219.
- (63) Kojaoghlanian T, Flomenberg P, Horwitz MS. The impact of adenovirus infection on the immunocompromised host. *Rev Med Virol* 2003 May-Jun;13(3):155-171.
- (64) Kojima T, Hashimoto Y, Watanabe Y, Kagawa S, Uno F, Kuroda S, et al. A simple biological imaging system for detecting viable human circulating tumor cells. *J Clin Invest* 2009 Oct;119(10):3172-3181.
- (65) Lafferty-Whyte K, Cairney CJ, Will MB, Serakinci N, Daidone MG, Zaffaroni N, et al. A gene expression signature classifying telomerase and ALT immortalization reveals an hTERT regulatory network and suggests a mesenchymal stem cell origin for ALT. *Oncogene* 2009 Oct 29;28(43):3765-3774.
- (66) Leary RJ, Kinde I, Diehl F, Schmidt K, Clouser C, Duncan C, et al. Development of personalized tumor biomarkers using massively parallel sequencing. *Sci Transl Med* 2010 Feb 24;2(20):20ra14.
- (67) Leon SA, Shapiro B, Sklaroff DM, Yaros MJ. Free DNA in the serum of cancer patients and the effect of therapy. *Cancer Res* 1977 Mar;37(3):646-650.
- (68) Levy MZ, Allsopp RC, Futcher AB, Greider CW, Harley CB. Telomere end-replication problem and cell aging. *J Mol Biol* 1992 Jun 20;225(4):951-960.
- (69) Lo, Y.M., Chan, K.C., Sun, H., Chen, E.Z., Jiang, P., Lun, F.M., Zheng, Y.W., Leung, T.Y., Lau, T.K., Cantor, C.R., et al. Maternal plasma DNA sequencing reveals the genome-wide genetic and mutational profile of the fetus. *Science Translational Medicine* 2010(2):61-91.
- (70) Lo YM, Zhang J, Leung TN, Lau TK, Chang AM, Hjelm NM. Rapid clearance of fetal DNA from maternal plasma. *Am J Hum Genet* 1999 Jan;64(1):218-224.
- (71) Ludwig JA, Weinstein JN. Biomarkers in cancer staging, prognosis and treatment selection. *Nat Rev Cancer* 2005 Nov;5(11):845-856.
- (72) Lui YY, Chik KW, Chiu RW, Ho CY, Lam CW, Lo YM. Predominant hematopoietic origin of cell-free DNA in plasma and serum after sex-mismatched bone marrow transplantation. *Clin Chem* 2002 Mar;48(3):421-427.
- (73) Lui YY, Woo KS, Wang AY, Yeung CK, Li PK, Chau E, et al. Origin of plasma cell-free DNA after solid organ transplantation. *Clin Chem* 2003 Mar;49(3):495-496.
- (74) Lun FM, Tsui NB, Chan KC, Leung TY, Lau TK, Charoenkwan P, et al. Noninvasive prenatal diagnosis of monogenic diseases by digital size selection and relative mutation dosage on DNA in maternal plasma. *Proc Natl Acad Sci U S A* 2008 Dec 16;105(50):19920-19925.
- (75) Maheswaran S, Haber DA. Circulating tumor cells: a window into cancer biology and metastasis. *Curr Opin Genet Dev* 2010 Feb;20(1):96-99.
- (76) Maida Y, Kyo S, Sakaguchi J, Mizumoto Y, Hashimoto M, Mori N, et al. Diagnostic potential and limitation of imaging cancer cells in cytological samples using telomerase-specific replicative adenovirus. *Int J Oncol* 2009 Jun;34(6):1549-1556.

- (77) Makarov VL, Hirose Y, Langmore JP. Long G tails at both ends of human chromosomes suggest a C strand degradation mechanism for telomere shortening. *Cell* 1997 Mar 7;88(5):657-666.
- (78) Maser RS, DePinho RA. Connecting chromosomes, crisis, and cancer. *Science* 2002 Jul 26;297(5581):565-569.
- (79) Masutomi K, Yu EY, Khurts S, Ben-Porath I, Currier JL, Metz GB, et al. Telomerase maintains telomere structure in normal human cells. *Cell* 2003 Jul 25;114(2):241-253.
- (80) Mathis JM, Stoff-Khalili MA, Curiel DT. Oncolytic adenoviruses - selective retargeting to tumor cells. *Oncogene* 2005 Nov 21;24(52):7775-7791.
- (81) McBride DJ, Orpana AK, Sotiriou C, Joensuu H, Stephens PJ, Mudie LJ, et al. Use of cancer-specific genomic rearrangements to quantify disease burden in plasma from patients with solid tumors. *Genes Chromosomes Cancer* 2010 Nov;49(11):1062-1069.
- (82) McClintock B. The Stability of Broken Ends of Chromosomes in Zea Mays. *Genetics* 1941 Mar;26(2):234-282.
- (83) McClintock B. The Production of Homozygous Deficient Tissues with Mutant Characteristics by Means of the Aberrant Mitotic Behavior of Ring-Shaped Chromosomes. *Genetics* 1938 Jul;23(4):315-376.
- (84) Misale S, Yaeger R, Hobor S, Scala E, Janakiraman M, Liska D, et al. Emergence of KRAS mutations and acquired resistance to anti-EGFR therapy in colorectal cancer. *Nature* 2012 Jun 28;486(7404):532-536.
- (85) Modrow S, Falke D, Truyen U. *Molekulare Virologie*. 2.th ed. Berlin: Spektrum Akademischer Verlag Heidelberg; 2003.
- (86) Moyzis RK, Buckingham JM, Cram LS, Dani M, Deaven LL, Jones MD, et al. A highly conserved repetitive DNA sequence, (TTAGGG)_n, present at the telomeres of human chromosomes. *Proc Natl Acad Sci U S A* 1988 Sep;85(18):6622-6626.
- (87) Muller I, Beeger C, Alix-Panabieres C, Rebillard X, Pantel K, Schwarzenbach H. Identification of loss of heterozygosity on circulating free DNA in peripheral blood of prostate cancer patients: potential and technical improvements. *Clin Chem* 2008 Apr;54(4):688-696.
- (88) Nagata S. DNA degradation in development and programmed cell death. *Annu Rev Immunol* 2005;23:853-875.
- (89) Nagrath S, Sequist LV, Maheswaran S, Bell DW, Irimia D, Utkus L, et al. Isolation of rare circulating tumour cells in cancer patients by microchip technology. *Nature* 2007 Dec 20;450(7173):1235-1239.
- (90) Nawroz H, Koch W, Anker P, Stroun M, Sidransky D. Microsatellite alterations in serum DNA of head and neck cancer patients. *Nat Med* 1996 Sep;2(9):1035-1037.
- (91) O'Flaherty JD, Gray S, Richard D, Fennell D, O'Leary JJ, Blackhall FH, et al. Circulating tumour cells, their role in metastasis and their clinical utility in lung cancer. *Lung Cancer* 2012 Apr;76(1):19-25.
- (92) Ohki R, Tsurimoto T, Ishikawa F. In vitro reconstitution of the end replication problem. *Mol Cell Biol* 2001 Sep;21(17):5753-5766.
- (93) Pantel K, Brakenhoff RH. Dissecting the metastatic cascade. *Nat Rev Cancer* 2004 Jun;4(6):448-456.
- (94) Partsalis T, Chan LY, Hurworth M, Willers C, Pavlos N, Kumta N, et al. Evidence of circulating donor genetic material in bone allotransplantation. *Int J Mol Med* 2006 Jun;17(6):1151-1155.
- (95) Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* 2001 May 1;29(9):e45.

- (96) Pinzani P, Salvianti F, Cascella R, Massi D, De Giorgi V, Pazzagli M, et al. Allele specific Taqman-based real-time PCR assay to quantify circulating BRAFV600E mutated DNA in plasma of melanoma patients. *Clin Chim Acta* 2010 Sep 6;411(17-18):1319-1324.
- (97) Pinzani P, Salvianti F, Pazzagli M, Orlando C. Circulating nucleic acids in cancer and pregnancy. *Methods* 2010 Apr;50(4):302-307.
- (98) Riethdorf S, Fritsche H, Muller V, Rau T, Schindlbeck C, Rack B, et al. Detection of circulating tumor cells in peripheral blood of patients with metastatic breast cancer: a validation study of the CellSearch system. *Clin Cancer Res* 2007 Feb 1;13(3):920-928.
- (99) Robbins PD, Ghivizzani SC. Viral vectors for gene therapy. *Pharmacol Ther* 1998 Oct;80(1):35-47.
- (100) Roychowdhury S, Iyer MK, Robinson DR, Lonigro RJ, Wu YM, Cao X, et al. Personalized oncology through integrative high-throughput sequencing: a pilot study. *Sci Transl Med* 2011 Nov 30;3(111):111ra121.
- (101) Russell WC. Adenoviruses: update on structure and function. *J Gen Virol* 2009 Jan;90(Pt 1):1-20.
- (102) Russell WC. Update on adenovirus and its vectors. *J Gen Virol* 2000 Nov;81(Pt 11):2573-2604.
- (103) Sanchez-Cespedes M, Esteller M, Wu L, Nawroz-Danish H, Yoo GH, Koch WM, et al. Gene promoter hypermethylation in tumors and serum of head and neck cancer patients. *Cancer Res* 2000 Feb 15;60(4):892-895.
- (104) Scher HI, Jia X, de Bono JS, Fleisher M, Pienta KJ, Raghavan D, et al. Circulating tumour cells as prognostic markers in progressive, castration-resistant prostate cancer: a reanalysis of IMMC38 trial data. *Lancet Oncol* 2009 Mar;10(3):233-239.
- (105) Schwarzenbach H, Alix-Panabieres C, Muller I, Letang N, Vendrell JP, Rebillard X, et al. Cell-free tumor DNA in blood plasma as a marker for circulating tumor cells in prostate cancer. *Clin Cancer Res* 2009 Feb 1;15(3):1032-1038.
- (106) Schwarzenbach H, Hoon DS, Pantel K. Cell-free nucleic acids as biomarkers in cancer patients. *Nat Rev Cancer* 2011 Jun;11(6):426-437.
- (107) Sedivy JM. Can ends justify the means?: telomeres and the mechanisms of replicative senescence and immortalization in mammalian cells. *Proc Natl Acad Sci U S A* 1998 Aug 4;95(16):9078-9081.
- (108) Sharma GG, Gupta A, Wang H, Scherthan H, Dhar S, Gandhi V, et al. hTERT associates with human telomeres and enhances genomic stability and DNA repair. *Oncogene* 2003 Jan 9;22(1):131-146.
- (109) Shay JW, Roninson IB. Hallmarks of senescence in carcinogenesis and cancer therapy. *Oncogene* 2004 Apr 12;23(16):2919-2933.
- (110) Shay JW, Wright WE. Senescence and immortalization: role of telomeres and telomerase. *Carcinogenesis* 2005 May;26(5):867-874.
- (111) Sidransky D. Emerging molecular markers of cancer. *Nat Rev Cancer* 2002 Mar;2(3):210-219.
- (112) Sleijfer S, Gratama JW, Sieuwerts AM, Kraan J, Martens JW, Foekens JA. Circulating tumour cell detection on its way to routine diagnostic implementation? *Eur J Cancer* 2007 Dec;43(18):2645-2650.
- (113) Smirnov DA, Zweitzig DR, Foulk BW, Miller MC, Doyle GV, Pienta KJ, et al. Global gene expression profiling of circulating tumor cells. *Cancer Res* 2005 Jun 15;65(12):4993-4997.
- (114) Smogorzewska A, de Lange T. Regulation of telomerase by telomeric proteins. *Annu Rev Biochem* 2004;73:177-208.
- (115) Song JS. Activity of the human telomerase catalytic subunit (hTERT) gene promoter could be increased by

the SV40 enhancer. *Biosci Biotechnol Biochem* 2004 Aug;68(8):1634-1639.

(116) Stott SL, Hsu CH, Tsukrov DI, Yu M, Miyamoto DT, Waltman BA, et al. Isolation of circulating tumor cells using a microvortex-generating herringbone-chip. *Proc Natl Acad Sci U S A* 2010 Oct 26;107(43):18392-18397.

(117) Stroun M, Anker P, Maurice P, Lyautey J, Lederrey C, Beljanski M. Neoplastic characteristics of the DNA found in the plasma of cancer patients. *Oncology* 1989;46(5):318-322.

(118) Sun YF, Yang XR, Zhou J, Qiu SJ, Fan J, Xu Y. Circulating tumor cells: advances in detection methods, biological issues, and clinical relevance. *J Cancer Res Clin Oncol* 2011 Aug;137(8):1151-1173.

(119) Takakura M, Kyo S, Kanaya T, Hirano H, Takeda J, Yutsudo M, et al. Cloning of human telomerase catalytic subunit (hTERT) gene promoter and identification of proximal core promoter sequences essential for transcriptional activation in immortalized and cancer cells. *Cancer Res* 1999 Feb 1;59(3):551-557.

(120) Talasz AH, Powell AA, Huber DE, Berbee JG, Roh KH, Yu W, 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* 2009 Mar 10;106(10):3970-3975.

(121) Tauber B, Dobner T. Molecular regulation and biological function of adenovirus early genes: the E4 ORFs. *Gene* 2001 Oct 31;278(1-2):1-23.

(122) Ulaner GA, Hu JF, Vu TH, Giudice LC, Hoffman AR. Telomerase activity in human development is regulated by human telomerase reverse transcriptase (hTERT) transcription and by alternate splicing of hTERT transcripts. *Cancer Res* 1998 Sep 15;58(18):4168-4172.

(123) Umeoka T, Kawashima T, Kagawa S, Teraishi F, Taki M, Nishizaki M, et al. Visualization of intrathoracically disseminated solid tumors in mice with optical imaging by telomerase-specific amplification of a transferred green fluorescent protein gene. *Cancer Res* 2004 Sep 1;64(17):6259-6265.

(124) Valastyan S, Weinberg RA. Tumor metastasis: molecular insights and evolving paradigms. *Cell* 2011 Oct 14;147(2):275-292.

(125) van der Vaart M, Pretorius PJ. Circulating DNA. Its origin and fluctuation. *Ann N Y Acad Sci* 2008 Aug;1137:18-26.

(126) Walther A, Johnstone E, Swanton C, Midgley R, Tomlinson I, Kerr D. Genetic prognostic and predictive markers in colorectal cancer. *Nat Rev Cancer* 2009 Jul;9(7):489-499.

(127) Wirth T, Zender L, Schulte B, Mundt B, Plentz R, Rudolph KL, et al. A telomerase-dependent conditionally replicating adenovirus for selective treatment of cancer. *Cancer Res* 2003 Jun 15;63(12):3181-3188.

(128) Wong JM, Collins K. Telomere maintenance and disease. *Lancet* 2003 Sep 20;362(9388):983-988.

(129) Wood LD, Parsons DW, Jones S, Lin J, Sjoblom T, Leary RJ, et al. The genomic landscapes of human breast and colorectal cancers. *Science* 2007 Nov 16;318(5853):1108-1113.

(130) Wu KJ, Grandori C, Amacker M, Simon-Vermot N, Polack A, Lingner J, et al. Direct activation of TERT transcription by c-MYC. *Nat Genet* 1999 Feb;21(2):220-224.

(131) Yin J, Fu P, Wang X, Jin G, Xiu R. Tumor-specific expression detected with the use of an expression vector driven by human telomerase reverse transcriptase gene promoter. *J Lab Clin Med* 2004 Dec;144(6):302-306.

(132) Yu M, Stott S, Toner M, Maheswaran S, Haber DA. Circulating tumor cells: approaches to isolation and characterization. *J Cell Biol* 2011 Feb 7;192(3):373-382.

(133) Yung TK, Chan KC, Mok TS, Tong J, To KF, Lo YM. Single-molecule detection of epidermal growth factor receptor mutations in plasma by microfluidics digital PCR in non-small cell lung cancer patients. Clin Cancer Res 2009 Mar 15;15(6):2076-2084.

13. Supplementary

#	sex	age (years)	Site primary	Diagnosis of primary tumor = Dukes' stage (T stage, N stage (no. Positive Lnn/no. Examined Lnn), M stage, grade)*	Diameter of primary lesion (cm)	KRAS (primary tumor)	Interval between diagnosis of primary and first metastasis (months)	Sites of metastases	Largest diameter of metastasis(cm)	Interval between diagnosis of primary and blood collection (months)	Last therapy	Interval between blood collection and last therapy (days)	Dukes' stage**	CEA** (ref.:0-5)	CA-19-9** (ref.: 0-37)
---	-----	----------------	-----------------	--	---------------------------------------	----------------------------	---	------------------------	---------------------------------------	---	-----------------	---	-------------------	---------------------	---------------------------

Supplementary Table1: Summary of clinical characteristics of the 32 colorectal carcinoma (CRC) patients

1	M	69	cecum	C (pT2 N2 (5/26) M0, G3)	6.4	G12V	Ø	none	Ø	2	2 cycles FOLFOX	12	C	3.4	21.40
2	F	48	colon sigmoideum	D (pT3 N2 (7/20), M1, G3)	6	ND	0	Liver	2	1	Surgery, no chemotherapy	45	D	2.5	8.30
3	M	50	rectum	C (pT3c N1(3/25) M0, G2)	4.5	WT	8	Liver	2.8	9	11 cycles FOLFOX	16	D	7.5	<2
6	M	60	cecum	D (pT3 N1(1/6), M1, G3)	NA	G12V	0	Liver, abdominal lymphnodes, peritoneal carcinomatosis, bone, brain	6	8	6 cycles Oxaliplatin	48	D	4315	370448.00
7	F	45	rectum	C (pT3 N1(3/20), M0, G3)	3	G13D	14	Lung	2	28	12 cycles FOLFOX Folfiri	313	D	16.7	74.50
9	M	70	rectum	D (only biopsy, therefore no data about T, N, G stage available), M1	NA	WT	0	Liver, lymphnodes, peritoneal carcinomatosis, bone	1.5	1	Surgery, no chemotherapy	34	D	39.4	2312.00
10	M	76	colon transversum	D (only biopsy, therefore no data about T, N, G stage available), M1		G12D	36	Liver	9.2	47	13 cycles FOLFOX	14	D	1935	7517.70
11	M	76	rectum	C (pT3 N2(10/22) M0, G1)	4.7	G12V	13	Liver, lung	3.5	63	8 cycles Irinotecan	99	D	4.2	25.40
12	M	75	cecum	D (pT4 N2(27/37)M1, G2)	8.5	G12A	36	Lung, abdominal lymphnodes	5.4	47	16 cycles FOLFIRI	95	D	3.4	16.90
14	M	69	colon sigmoideum	C (pT3 N0 (0/29), M0, G2)	4.5	ND	13	Lung, thoracal lymphnodes	3.2	20	6 cycles FOLFOX	89	D	8.9	30.60
15	F	74	cecum	D (pT3, N1(1/14),M1, G1)	4.5	G12D	1	Liver, lung	5	16	6 cycles FOLFIRI	22	D	106.4	279.80
16	M	71	colon descendens	D (pT3bN2(40/41)MX)	1.3	G12D	13	abdominal lymphnodes	NA	16	4 cycles Capecitabine	15	D	5	50.80
17	F	72	colon sigmoideum	D (pT4, N1 (3/5), M1, G2)	NA	WT	0	Liver, lung, lymphnodes, peritoneal	3.5	39	7 cycles FOLFOX	403	D	87.1	99.10

							carcinomatosis								
18	M	54	recto-sigmoideum	D (only biopsy, therefore no data about T, N, G stage available), M1	NA	WT	0	Liver	5.9	1	Ø	0	D	856.6	33.50
19	M	67	colon sigmoideum	D (pT3 N2(5/11), M1, G2)	7	G13D	0	Liver	1.2	15	12 cycles FOLFOX	216	D	2.2	<2
20	M	81	flexura coli sinistra	D (only biopsy, therefore no data about T, N, G stage available), M1	NA	WT	0	Liver	7	14	12 cycles Irinotecan	28	D	194.9	9192.70
21	M	76	rectum	C (pT3 N1(2/13), M0, G2)	NA	G12V	10	Liver, lung, bone	2	61	4 cycles Capecitabine	688	D	17.4	20.30
22	M	62	flexura coli sinistra	D (pT3 N2 (7/13) M1, G2)	0.3	WT	0	Liver, peritoneal carcinomatosis, adrenal gland	14.5	2	Surgery, no chemotherapy	2	D	1.6	<2
23	M	71	rectum	D (only biopsy, therefore no data about T, N, G stage available), M1	NA	WT	0	Liver, primary tumor in situ	10.6	1	Ø	0	D	10.5	26.20
24	M	69	colon transversum	D (pT3 N1 (2/17) M1, G2)	5.5	WT	18	Liver, peritoneal carcinomatosis	3.3	41	10 cycles FOLFIRI/ Bevacizumab	185	D	2.9	11.30
25	M	74	colon sigmoideum	D (only biopsy, therefore no data about T, N, G stage available), M1	NA	G12D	0	Liver, primary tumor in situ	6	8	7 cycles FOLFOX	67	D	5154.2	11881.20
26	F	70	colon transversum	C (pT3 N1 (2/40)M0, G2)	2.3	WT	31	Liver, bone, spleen	3.5	34	12 cycles Irinotecan / Cetuximab	34	D	95.4	64.70
27	F	66	sigma	D (pT4 N1(3/20), M1, G2)	4.5	WT	0	Liver	3.5	29	9 cycles FOLFIRI	110	D	2439.2	162.10
28	M	72	rectum	C (pT3 N0(0/19), M0, G2)	3.1	WT	134	Liver, lung, adrenal gland	2.5	151	10 cycles Irinotecan	138	D	227.6	168.70

29	F	76	rectum	C (pT3 N0(0/21), M0, G2)	4.3	WT	11	Liver	2.2	38	10 cycles FOLFOX	21	D	38.8	9.70
30	M	71	colon ascendens	C (pT3 N1(2/20) M0, G2)	4	WT	7	Liver	2.6	27	5 cycles FOLFOX	20	D	3689.5	18.70
32	F	67		D (pT4 N2 (5/16) M1, G2)			0	Peritoneal carcinomatosis		10	8 cycles FOLFOX	69	D	5.1	70.10
33	F	75	colon ascendens	C (pT4 N0 (0/12) M1, G2)	4	WT	0	Liver, peritoneal carcinomatosis, abdominal wall metastasis	9.2	9	8 cycles Capecitabine	43	D	1303.3	4.10
34	M	77	rectum	C (pT3 N2 (5/8) M0, G2)	3	G12D	35	Liver, lung	7	62	3 cycles Tegafur-Uracil (07/2010-10/2010)	121	D	1075.1	456.50
35	M	78	sigma	D (pT4 N1 (3/10) M1, G2)	5	WT	0	Lung, bone, lymph nodes	1.5	94	4 cycles FOLFIRI (10/2010-12/2010)	104	D	1.1	11.90
37	M	67	sigma	D (pT4, N1 (2/15), M1, G2)	3	G13D	61	Peritoneal carcinomatosis		74	5 cycles FOLFIRI/Bevacizumab	277	D	10.2	208.10
38	M	63	colon ascendens	B (pT3, N0 (0/28), M0, G2)	5	G12D	44	Liver, lung, bone, lymph nodes	9	45	Surgery, no chemotherapy	0	D	2.6	19.20

*: at time of diagnosis; **: at time of blood collection

M: male

F: female

WT: wild type

y: years m: months d: days, ND: not done

NA: not available

Supplementary Table 2.

Ultra-deep pyrosequencing with sequencing reaction sizes of 119 bp, 168 bp, and 323 bp in patients without (#7, #11, #15, and #16) and with biphasic plasma-DNA distribution (#6, #10, #25, #38). For each sample the top line indicates the percentage of mutated DNA fragments (column “% mutated”) and the number of reads (column “reads”), whereas the bottom line indicates the same after whole-genome amplification (WGA). The column labeled “Mutation” lists the *KRAS* mutation found in the respective primary tumor.

Patient	Mutation	119 bp		168 bp		323 bp	
		% mutated	reads	% mutated	reads	% mutated	reads
#7	G13D	0	11249	0	22059	0	18643
	(c.38G>A)	0	9939	0	22059	0	0
#11	G12V	1.18	8082	0.31	9717	0	21010
	(c.35G>T)	0	7845	0	8498	0	3342
#15	G12D	0	12649	6.9	11254	0	0
	(c.35G>A)	0	13401	0	22527	0	29546
#16	G12D	0	9406	0	8553	0	6961
	(c.35G>A)	0.62	13271	0	11745	0.36	6049
#6	G12V	39.25	14114	42.72	8978	48.34	6492
	(c.35G>T)	23.42	5479	9.61	12894	19.74	2989
#10	G12D	33.06	189540	33.33	17608	27.24	22444
	(c.35G>A)	49.08	12270	28.42	21641	25.47	14447
#25	G12D	17.11	13171	12.54	12163	0	4169
	(c.35G>A)	13.23	12732	8.59	13956	0	10537
#38	G12D	37.47	11392	33.32	12778	27.37	21057
	(c.35G>A)	10.56	14596	4.56	14292	3.88	20450

Supplementary Table 3.

Sites of metastases and their association with the biphasic plasma DNA (bi-pDNA) distribution and number of CTCs (NA: not available).

			Number of CTCs					
Metastasis site	number of patients with bi-pDNA	number of patients without bi-pDNA	0	1	2-5	6-10	>10	NA
bone	4	2	1				4	1
liver	11	14	7	3	4	3	6	2
lung	1	10	4	3	2		1	1
peritoneal carcinomatosis	3	5		2	3		3	

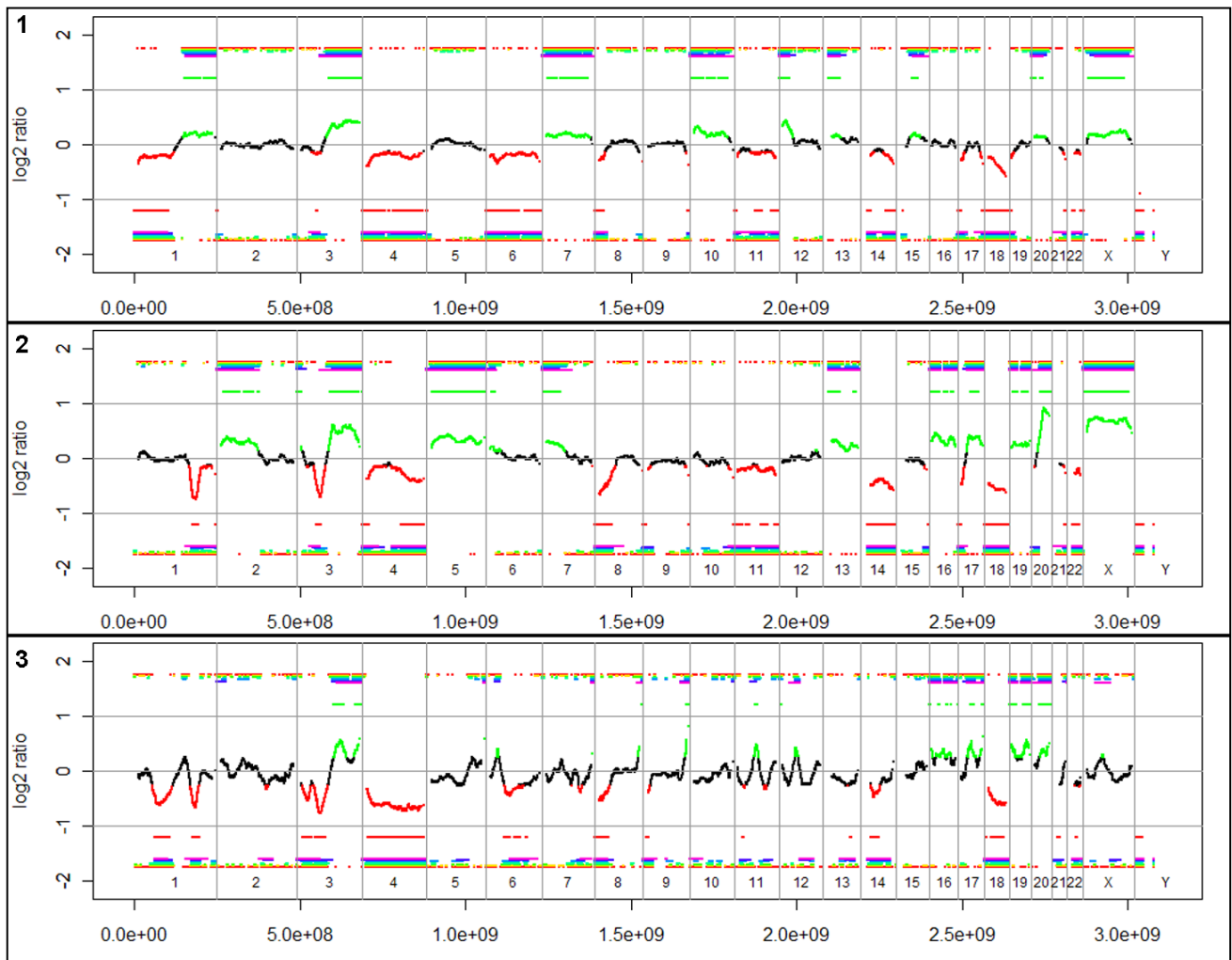
Supplementary Table 4.

Mapping of regions involved in artifacts after whole-genome amplification (WGA) of severely fragmented plasma DNA. The first column lists the chromosomal region, the second and third columns the exact borders of the respective region, the fourth column the size of the region (in bp), the fifth column the number of oligos on the array-platform involved in the artifact, and the column labeled “in/de” whether the artifact resulted in increased or decreased ratio values.

Chromosome	Start	End	Size (bp)	# Oligos	in/de
1p36.33-1p35.1	1:749625	1:33954225	33204600	810	in
1p34.3	1:36183989	1:38235282	2051293	54	in
1q21.2-1q22	1:149245378	1:155042676	5797298	262	in
2q24.2-2q31.1	2:162544755	2:170471439	7926684	174	de
2q31.1-2q32.3	2:182580451	2:196295211	13714760	204	de
2q37.3	2:240714768	2:242690037	1975269	50	in
4p16.3	4:950001	4:3550117	2600116	54	in
4q13.1-4q13.3	4:65962632	4:74827416	8864784	149	de
4q32.1	4:156021483	4:157480844	1459361	34	de
4q32.3-4q34.1	4:167206304	4:175832122	8625818	164	de
5q12.1-5q12.3	5:60977290	5:65345244	4367954	79	de
6p21.33-6p21.31	6:30422708	6:34897675	4474967	199	in
6p21.2	6:36759419	6:38701549	1942130	59	in
6p21.1	6:41227347	6:43038254	1810907	56	in
6q21-6q22.33	6:114397770	6:129094132	14696362	254	de
8q13.3-8q21.12	8:71939727	8:79170796	7231069	109	de
9q33.3-9q34.3	9:128026884	9:140128743	12101859	358	in
11p15.5	11:1	11:2246648	2246648	49	in
11q12.2-11q13.4	11:60699074	11:72872679	12173605	439	in
12q21.31-12q22	12:81394905	12:94217898	12822993	179	de
12q22-12q23.1	12:95315313	12:96618701	1303388	29	de
12q24.23- 12q24.31	12:119017413	12:124390885	5373472	174	in

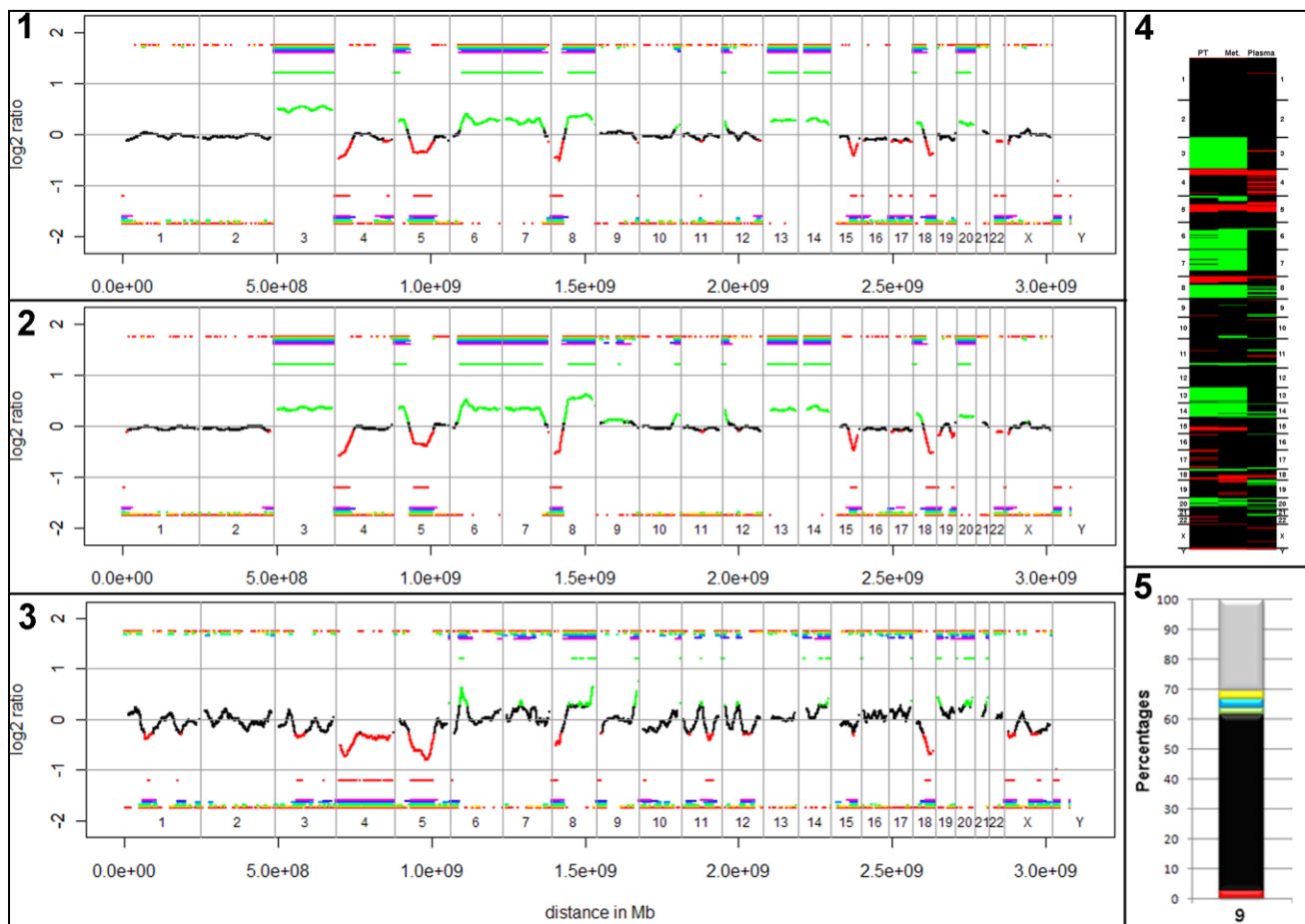
12q24.33	12:130299366	12:132231339	1931973	59	in
13q14.3-13q21.33	13:52406229	13:72583724	20177495	279	de
13q22.3-13q32.1	13:78632731	13:96195296	17562565	279	de
14q32.32- 14q32.22	14:103527487	14:105467584	1940097	57	in
16p13.3	16:1	16:5856947	5856947	199	in
16p12.2-16p11.2	16:22221294	16:31412133	9190839	249	in
16q21-16q22.1	16:65151089	16:67433141	2282052	124	in
16q22.1-16q22.3	16:68534348	16:73524277	4989929	124	in
16q23.3-16q24.2	16:81990225	16:88610804	6620579	209	in
17p13.3-17p13.1	17:1	17:8970183	8970183	334	in
17p11.2-17q11.1	17:16186659	17:25018087	8831428	244	in
17q11.2	17:26473801	17:28216578	1742777	49	in
17q12-17q21.32	17:33805689	17:45980825	12175136	494	in
17q24.3-17q25.3	17:69712217	17:78623226	8911009	319	in
19p13.3-19p13.11	19:1	19:19854958	19854958	844	in
19q13.12-19q13.2	19:37823098	19:41577039	3753941	159	in
19q13.2-19q13.43	19:42864333	19:60859667	17995334	799	in
19q13.43	19:63251545	19:63784386	532841	34	in
20p13	20:1	20:3012700	3012700	74	in
20p11.21- 20q11.23	20:25213567	20:36325723	11112156	224	in
20q13.11-20q13.2	20:41756635	20:49881938	8125303	224	in
20q13.31- 20q13.32	20:55135683	20:57033197	1897514	49	in
20q13.33	20:59271072	20:62363629	3092557	107	in
21q11.2-21q21.1	21:14847723	21:17047504	2199781	49	de
21q22.2-21q22.3	21:42395372	21:46892353	4496981	169	in
22q11.1-22q11.23	22:15476913	22:23931833	8454920	219	in
22q12.1-22q13.33	22:27437979	22:49524240	22086261	629	in

Supplementary Figures



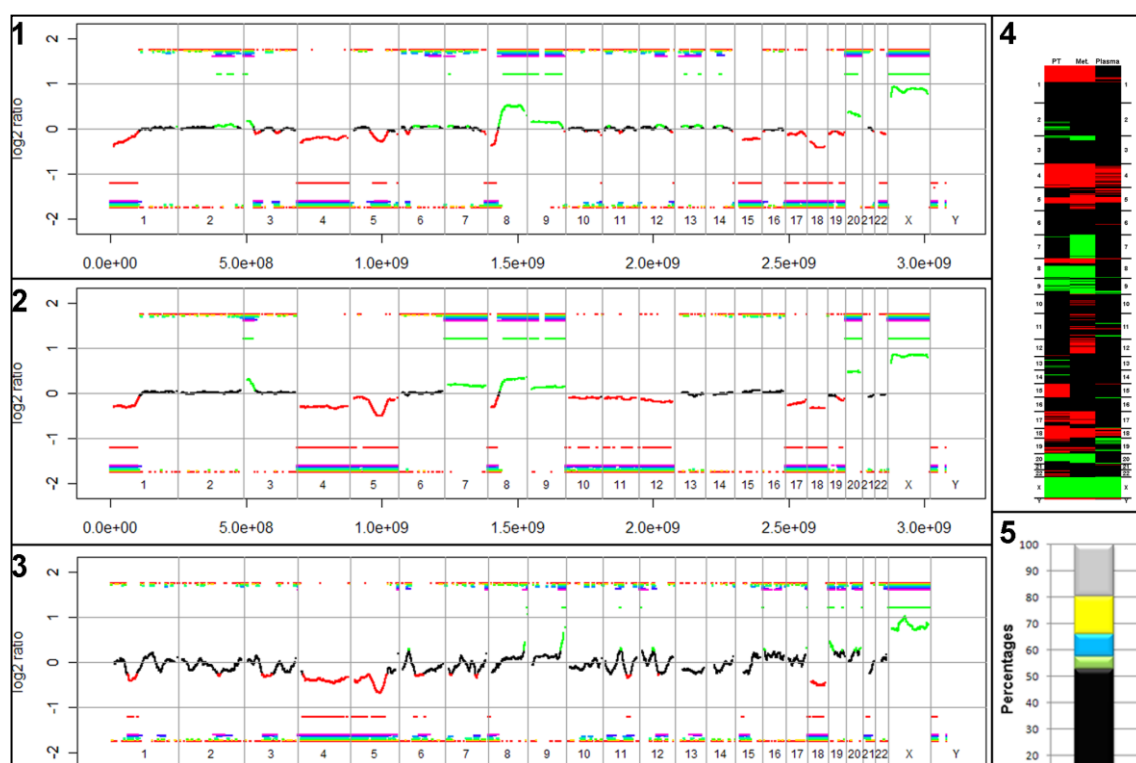
Suppl. Figure 1: Evaluation of available material for patient #6.

For all array-CGH profiles the multicolor bar codes at the top or bottom of the ratio profiles illustrate the results obtained during the iterative calculations with various window sizes, the single green and red bars summarize the regions which were gained or lost based on all calculations (Methods). Black parts in the profile represent balanced regions, lost regions appear in red, and gained regions in green. The respective heat maps show the copy number status (red: under-represented; black: balanced; green: over-represented). The bar charts display the percentages of chromosomal regions, which were commonly lost (red), balanced (black), or gained (green) in all analyzed samples; shared by metastasis and plasma DNA only (blue); shared by primary tumor and plasma DNA only (yellow); or unique to the plasma DNA (gray). (a) array-CGH profiles of the primary tumor (panel 1), metastasis (panel 2), and plasma DNA (panel 3) for patient #6.



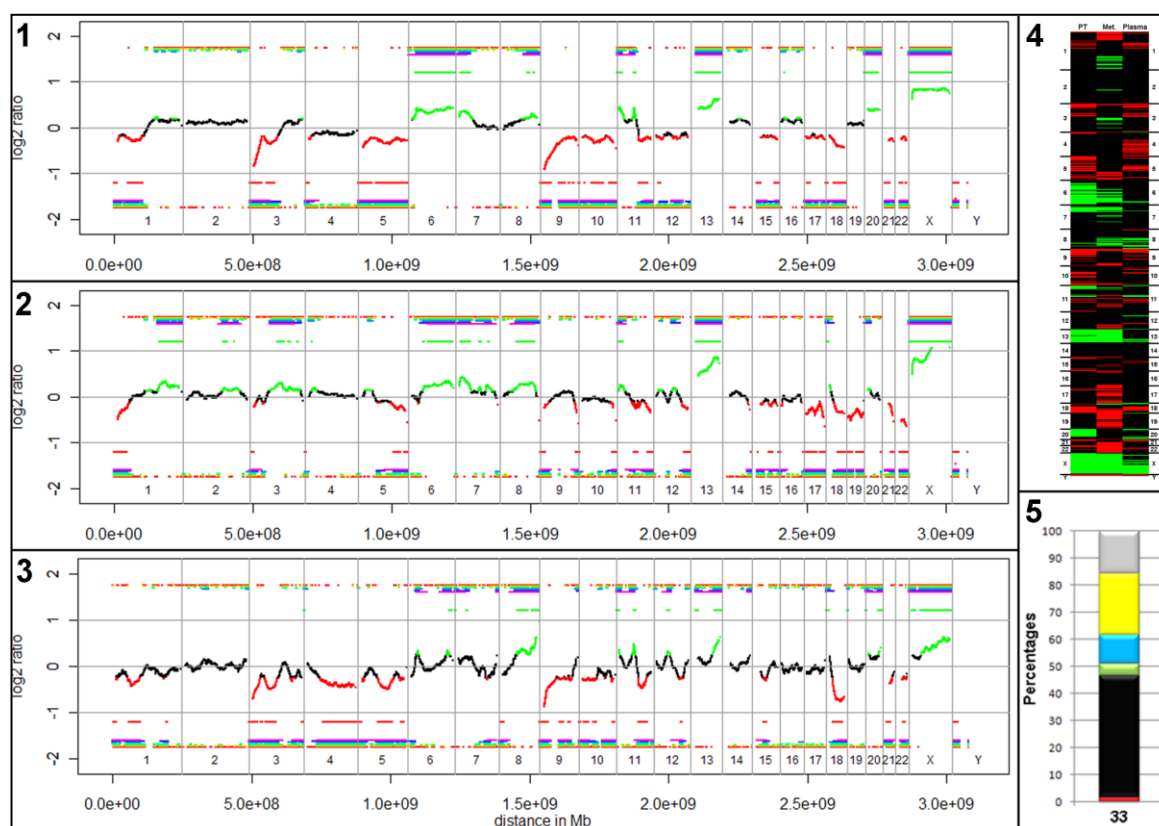
Suppl. Figure 2: Analysis of material from patient #9.

Array-CGH profiles of the primary tumor (panel 1), metastasis (peritoneal carcinomatosis; panel 2), plasma DNA (panel 3), heat map (panel 4), and bar chart (panel 5). For explanations regarding the color codes see legend of Supplementary Figure 1.



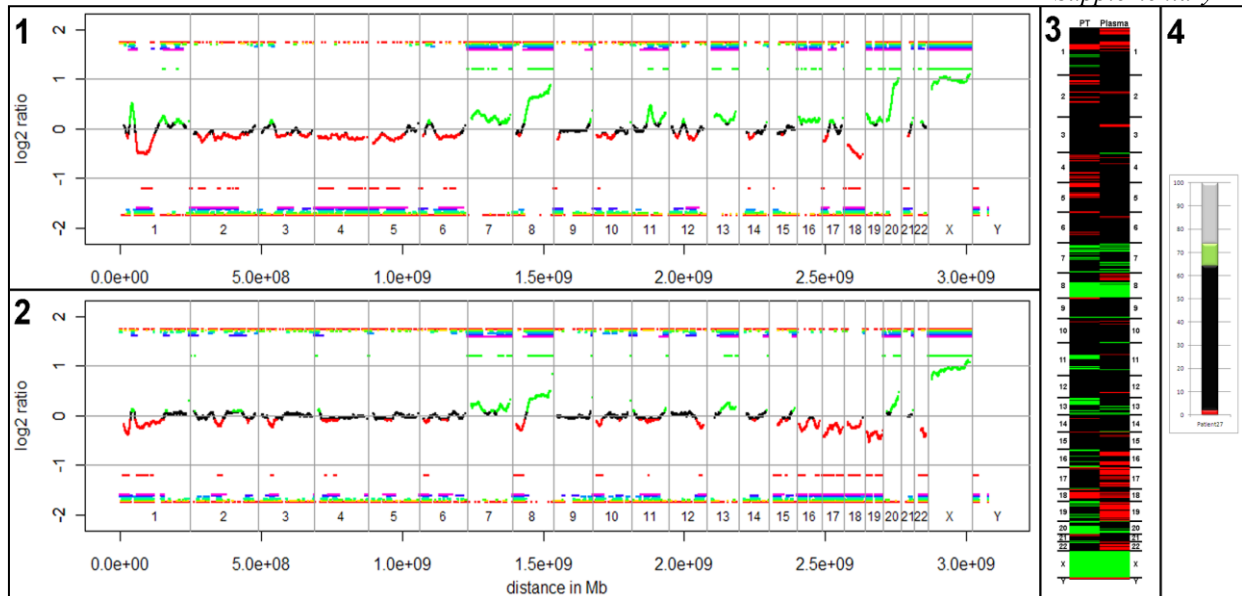
Suppl. Figure 3: Analysis of material from patient #26.

Array-CGH profiles of the primary tumor (panel 1), liver metastasis (panel 2), plasma DNA (panel 3), heat map (panel 4), and bar chart (panel 5). For explanations regarding the color codes see legend of Supplementary Figure 1.



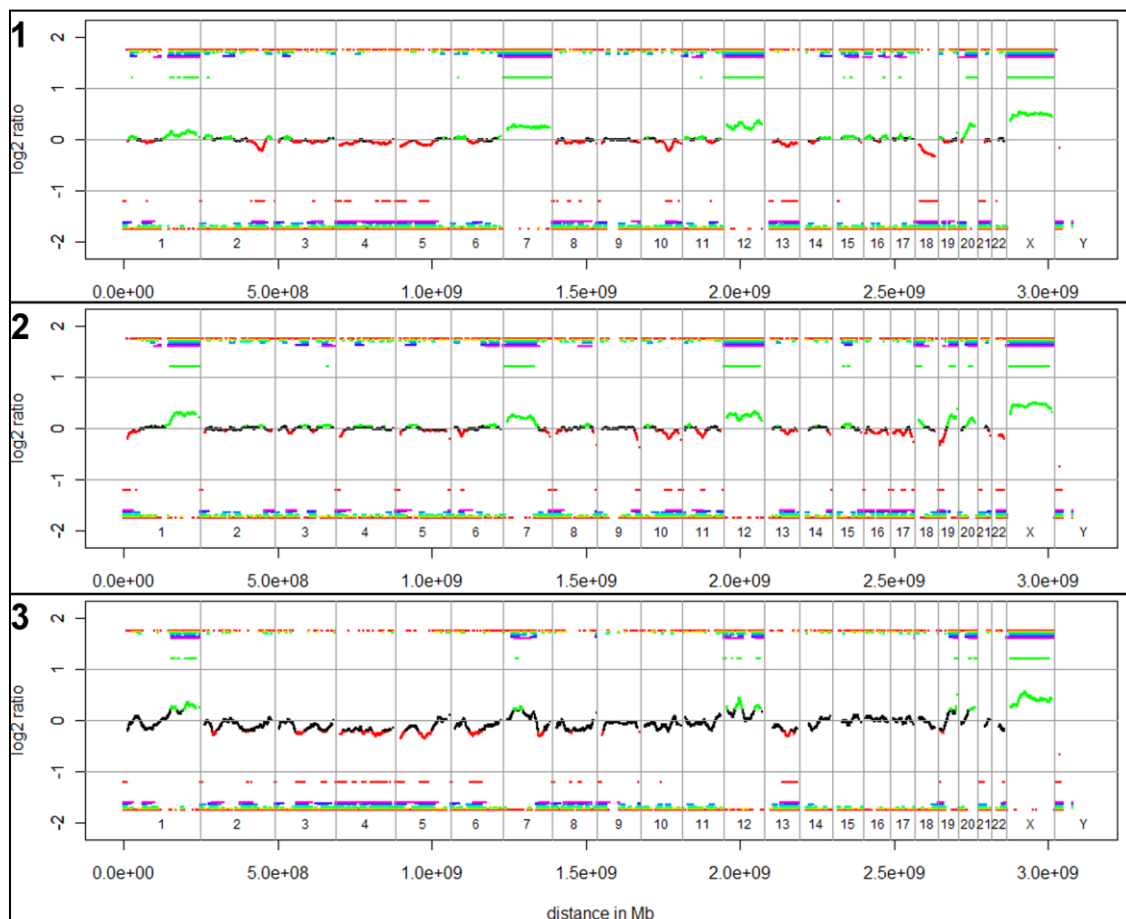
Suppl. Figure 4: Analysis of material from patient #33.

Array-CGH profiles of the primary tumor (panel 1), metastasis (liver; panel 2), plasma DNA (panel 3), heat map (panel 4), and bar chart (panel 5). For explanations regarding the color codes see legend of Supplementary Figure 1.



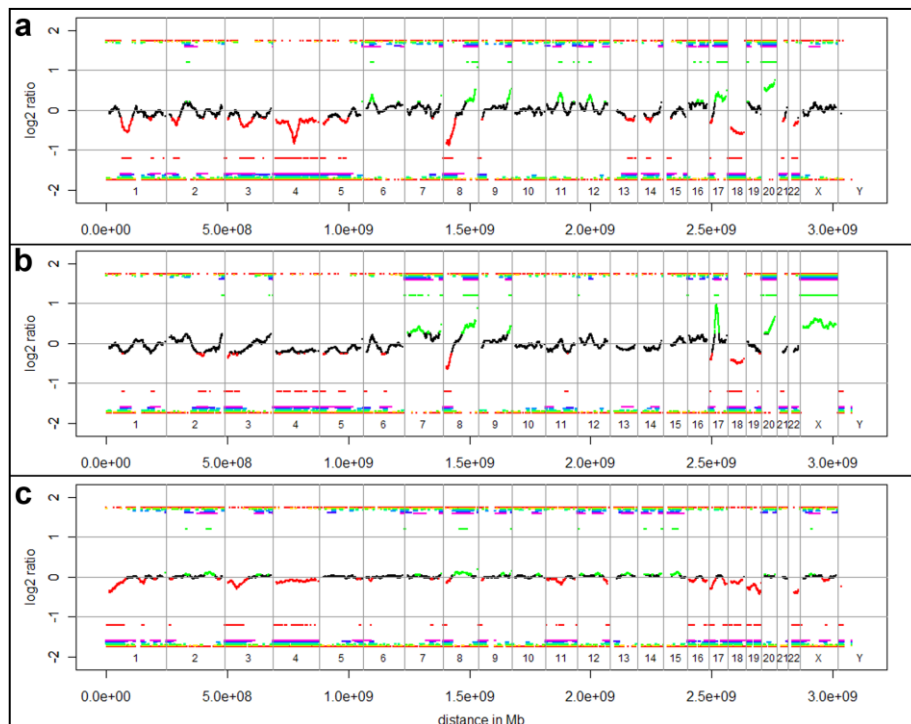
Suppl. Figure 5: Analysis of material from patient #27.

Analysis of primary tumor (panel 1), plasma DNA (panel 2), and heat map (panel 3). The bar chart (panel 4) indicates percentages of commonly lost (red), gained (green), and balanced (black) oligonucleotides. Gray indicates copy number changes observed only in the plasma DNA, but not in the primary tumor.

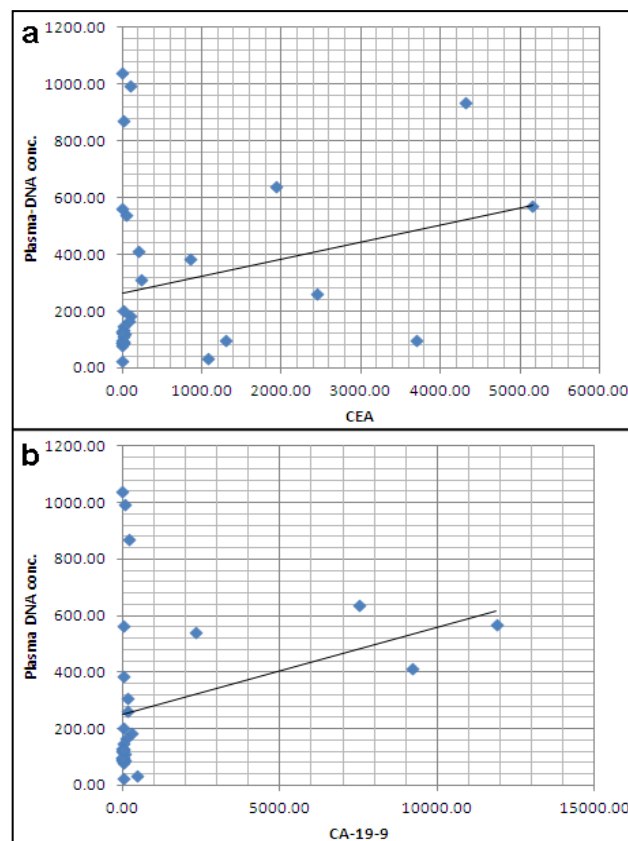


Suppl. Figure 6: Analysis of material from patient #38.

Array-CGH profiles of the primary tumor (panel 1), liver metastasis (panel 2), and plasma DNA (panel 3).



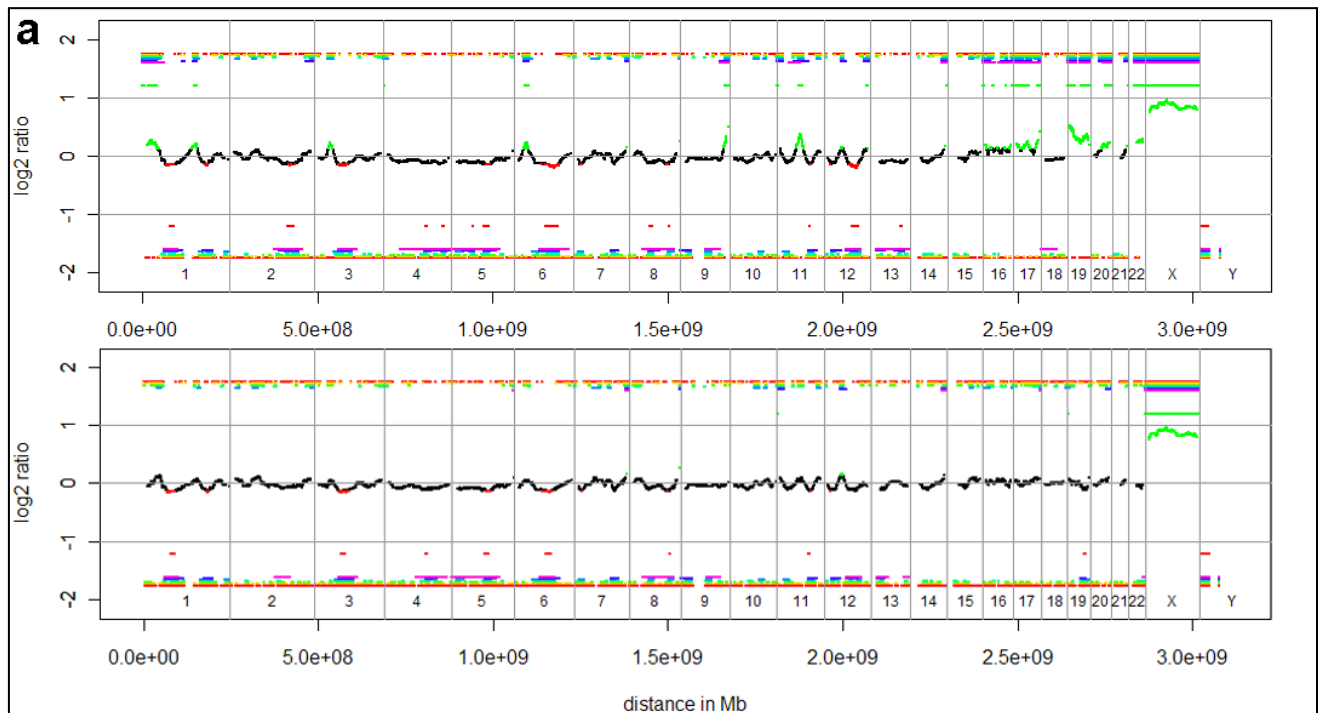
Suppl. Figure 7: Plasma DNA profiles from three patients for which no further tumor material was available for analysis, i.e. patient #10 (a); patient #20 (b); patient #25 (c).



Suppl. Figure 8: Correlation with tumor markers:

(a) CEA (slope of the linear regression line: 0.0601; correlation coefficient: 0.27).

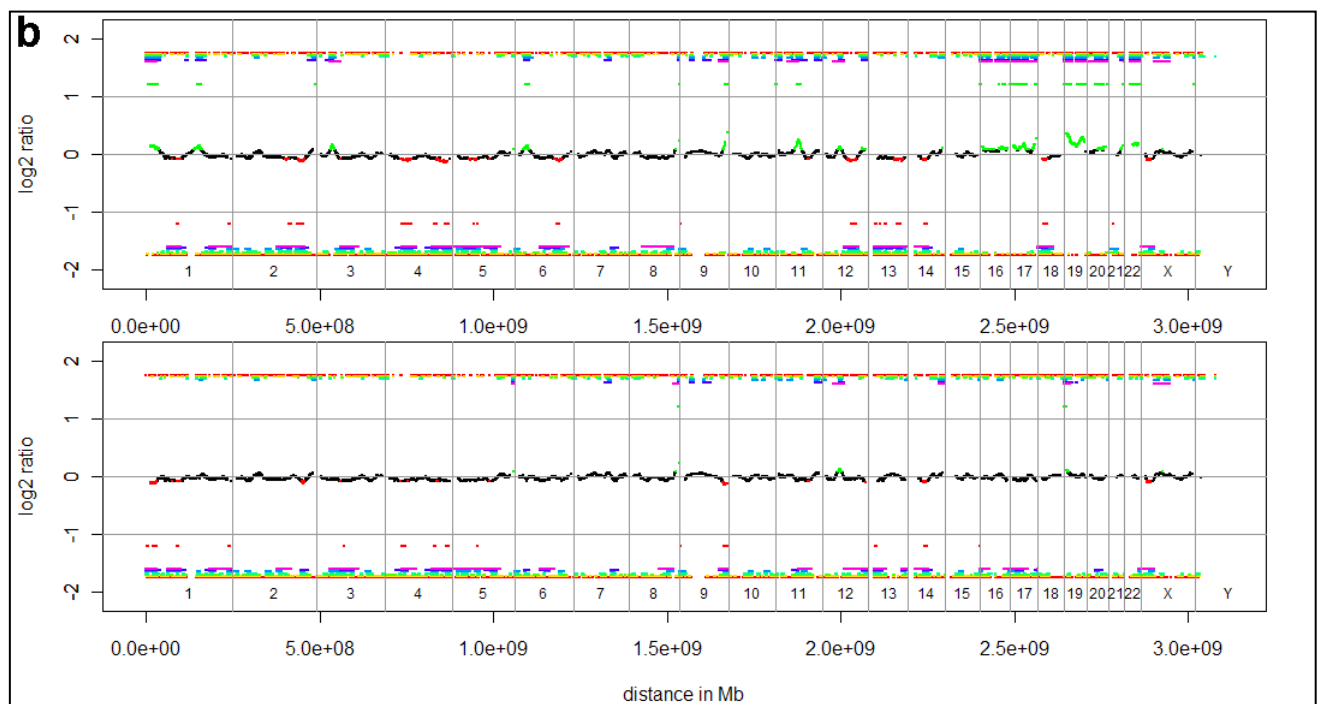
(b) CA19-9 (graph without patient 6, whose CA19-9 value was exceptionally high with 370,448; Slope of the linear regression line without patient 6: 0.0308; correlation coefficient: 0.315; with patient 6: slope: 0.0018; correlation coefficient: 0.394).



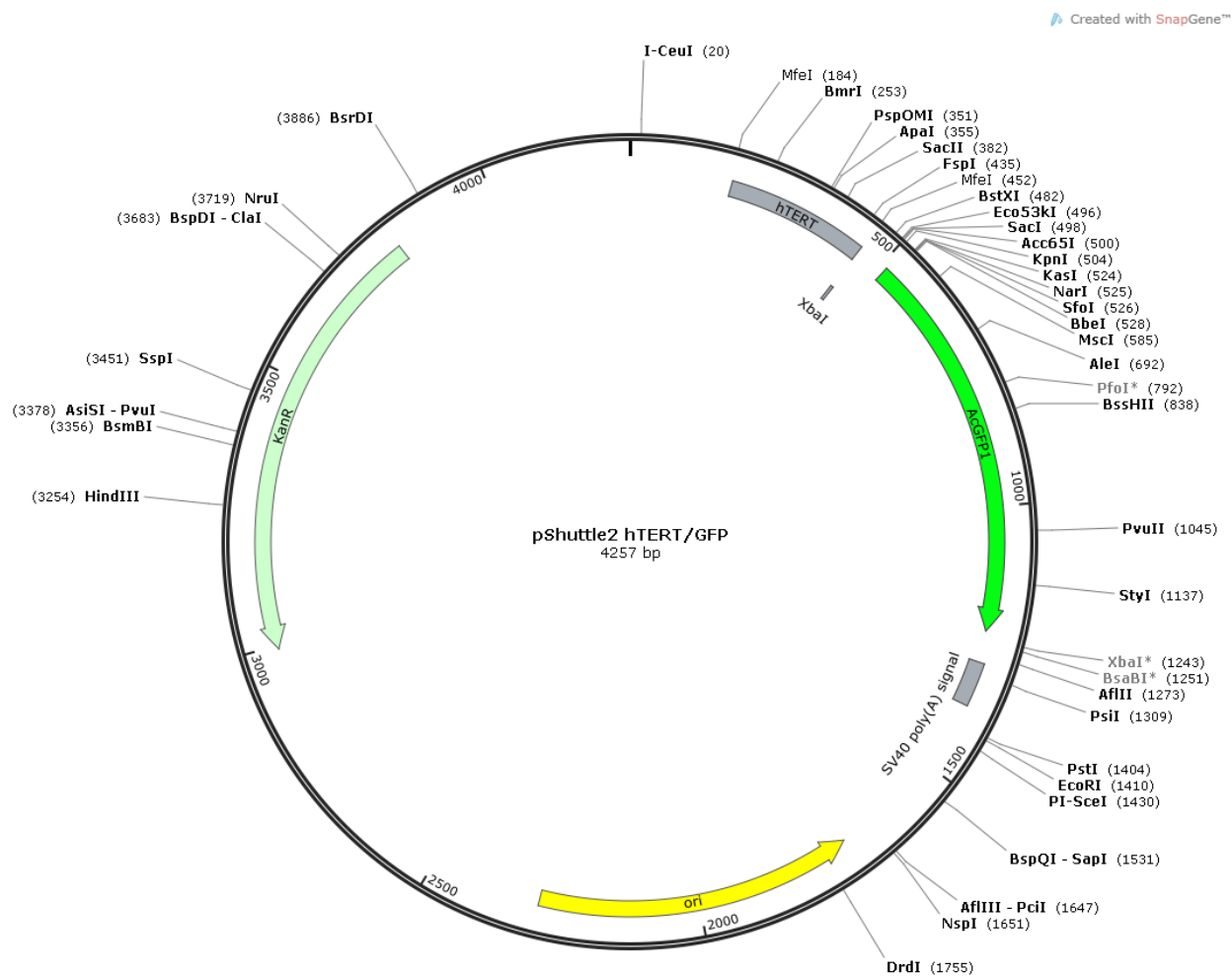
Suppl. Figure 9a: Array-CGH analysis after whole genome amplification products from plasma DNA.

(a) Plasma DNA analysis from female healthy donor F3 showing artifacts as listed in Supplementary Table 6 before (upper panel) and after correction (lower panel). Note that the overrepresentation of the X-chromosome (male reference DNA) is clearly visible because male reference DNA was used.

(b) As in (a), but with plasma DNA from male healthy donor M2 before (upper panel) and after (lower panel) artifact correction.



Suppl. Figure 9b: Array-CGH analysis after whole genome amplification products from plasma DNA.



Suppl. Figure 10: Vector map: pShuttle/hTERT/ GFP

Vektorkarte pShuttle2**Lablife**

(Source: https://www.lablife.org/g?a=seqa&id=vdb_g2.6UeXzl.c0l4s7tWLSixvj6Qdrjk-_sequence_95a3f841d050818c9d38580235f8255d831a99bb_10)

LOCUS pShuttle2 3974 bp DNA SYN 22-Oct-2012

DEFINITION pShuttle2

ACCESSION

KEYWORDS

SOURCE

ORGANISM other sequences; artificial sequences; vectors.

FEATURES Location/Qualifiers

source 1..3974

/organism="pShuttle2"

/mol_type="other

DNA"

misc_feature 338..625

/label="CAG_enhancer"

promoter 259..875

/label="CMV_immearyly_promoter"

misc_feature 792..812

/label="CMV_fwd_primer"

promoter 886..904

/label="T7_promoter"

misc_feature 1098..1117

/label="EBV_rev_primer"

rep_origin complement(1410..2029)

/label="pBR322_origin"

CDS complement(2710..3525)

/label="KanR2"

/gene="KanR2"

/note="ORF

frame

1"

/translation="MSHIQRETSCSRPRLNSNMDADLYGYKWARDNVGQSGATIYRLY

GKPDAPELFLKHGKGSVANDVTDEMVRNLNWLTEFMPLPTIKHFIRTPDDAWLLTTAIP

GKTAFQVLEEYPDSGENIVDALAVFLRRLHSIPVCNCPFNSDRVFLAQAQSRMNNGL

VDASDFDDERNGWPVEQVWKEMHKLLFPSPDSVVTHGDFSLDNLIFDEGKLIGCIDVG

RVGIADRYQDLAILWNCLGEFSPSLQKRLFQKYGIDNPDMNKLQFHLMLDEFF**

gene complement(2710..3525)

/label="KanR2"

/gene="KanR2"

promoter complement(3884..3912)

/label="AmpR_promoter"

ORIGIN

1 CGTAACTATA ACGGTCCTAA GGTAGCGAAA GCTCAGATCT CCCGATCCCC TATGGTGCAC
 61 TCTCAGTACA ATCTGCTCTG ATGCCGCATA GTTAAGCCAG TATCTGCTCC CTGCTTGTGT
 121 GTTGGAGGTC GCTGAGTAGT GCGCGAGCAA AATTTAAGCT ACAACAAGGC AAGGCTTGAC

MfeI

181 CGACATTGC ATGAAGAATC TGCTTAGGGT TAGGCGTTTT GCGCTGCTTC GCGATGTACG
 241 GGCCAGATAT ACGCGTTGAC ATTGATTATT GACTAGTTAT TAATAGTAAT CAATTACGGG
 301 GTCATTAGTT CATAGCCCAT ATATGGAGTT CCGCGTTACA TAACTTACGG TAAATGGCCC
 361 GCCTGGCTGA CCGCCCAACG ACCCCCGCCC ATTGACGTCA ATAATGACGT ATGTTCCCAT
 421 AGTAACGCCA ATAGGGACTT TCCATTGACG TCAATGGGTG GAGTATTTAC GGTAAACTGC
 481 CCACTTGGA GTACATCAAG TGTATCATAT GCCAAGTACG CCCCTATTG ACGTCAATGA
 541 CGGTAAATGG CCCGCCTGGC ATTATGCCCA GTACATGACC TTATGGGACT TTCCTACTTG
 601 GCAGTACATC TACGTATTAG TCATCGCTAT TACCATGGTG ATGCGGTTTT GGCAGTACAT
 661 CAATGGGCGT GGATAGCGGT TTGACTCACG GGGATTCCA AGTCTCCACC CCATTGACGT
 721 CAATGGGAGT TTGTTTTGGC ACCAAAATCA ACGGGACTTT CCAAATGTC GTAACAATC
 781 CGCCCCATTG ACGCAAATGG GCGGTAGGCG TGTACGGTGG GAGGTCTATA TAAGCAGAGC
 841 TCTCTGGCTA ACTAGAGAAC CCACTGCTTA CTGGCTTATC GAAATTAATA CGACTCACTA

XbaI

901 TAGGGAGACC CAAGCTGGCT AGCGTTTAAA CGGGCCCTCTAGACTCGAGC GGCCGCCACT

KpnIAflII

961 GTGCTGGATG ATCCGAGCTC GGTACCAAGC TTAAGTAAGT GACTAGAACT TGTATTATGC
 1021 AGCTTATAAT GGTTACAAAT AAAGCAATAG CATCACAAAT TTCACAAATA AAGCATTITT
 1081 TTCACTGCAT TCTAGTTGTG GTTTGTCCAA ACTCATCTGC AGATCTGAAT TCATCTATGT
 1141 CGGGTGCGGA GAAAGAGGTA ATGAAATGGC ATTATGGGTA TTATGGGTCT GCATTAATGA
 1201 ATCGGCCAAC GCGCGGGGAG AGGCGGTTTG CGTATTGGGC GCTCTTCCGC TTCCTCGCTC
 1261 ACTGACTCGC TCGCTCGGT CGTTCGGCTG CGGCGAGCGG TATCAGCTCA CTCAAAGGCG
 1321 GTAATACGGT TATCCACAGA ATCAGGGGAT AACGCAGGAA AGAACATGTG
 AGCAAAAGGC

1381 CAGCAAAAGG CCAGGAACCG TAAAAAGGCC GCGTTGCTGG CGTTTTTCCA TAGGCTCCGC
 1441 CCCCTGACG AGCATCACAA AAATCGACGC TCAAGTCAGA GTGGCGGAAA
 CCCGACAGGA

1501 CTATAAAGAT ACCAGGCGTT TCCCCTGGA AGCTCCCTCG TGCGCTCTCC TGTTCCGACC
 1561 CTGCCGCTTA CCGGATACCT GTCCGCTTT CTCCCTTCGG GAAGCGTGGC GCTTTCTCAA
 1621 TGCTCACGCT GTAGGTATCT CAGTTCGGTG TAGGTCGTTC GCTCCAAGCT GGGCTGTGTG
 1681 CACGAACCCC CCGTTCAGCC CGACCGCTGC GCCTTATCCG GTAACATCG TCTTGAGTCC
 1741 AACCCGGTAA GACACGACTT ATCGCCACTG GCAGCAGCCA CTGGTAACAG GATTAGCAGA
 1801 GCGAGGTATG TAGGCGGTGC TACAGAGTTC TTGAAGTGGT GGCCTAACTA CGGCTACACT
 1861 AGAAGGACAG TATTTGGTAT CTGCGCTCTG CTGAAGCCAG TTACCTTCGG AAAAAGAGTT
 1921 GGTAAGTCTT GATCCGGCAA ACAAACCACC GCTGGTAGCG GTGGTTTTTT TGTTTGCAAG
 1981 CAGCAGATTA CGCGCAGAAA AAAAGGATCT CAAGAAGATC CTTTGATCTT TTCTACGGGG

2041 TCTGACGCTC AGTGGAACGA AAACCTCACGT TAAGGGATTT TGGTCATGAG ATTATCAAAA
 2101 AGGATCTTCA CCTAGATCCT TTTGATCCTC CGGCGTTCAG CCTGTGCCAC AGCCGACAGG
 2161 ATGGTGACCA CCATTTGCCC CATATCACCG TCGGTACTGA TCCCGTCGTC AATAAACCGA
 2221 ACCGCTACAC CCTGAGCATC AAACCTCTTTT ATCAGTTGGA TCATGTCCGC GGTGTCCGCG
 2281 CCAAGACGGT CGAGCTTCTT CACCAGAATG ACATCACCTT CCTCCACCTT CATCCTCAGC
 2341 AAATCCAGCC CTCCCCGATC TGTTGAACTG CCGGATGCCT TGTCGGTAAA GATGCGGTTA
 2401 GCTTTTACCC CTGCATCTTT GAGCGCTGAG GTCTGCCTCG TGAAGAAGGT GTTGCTGACT
 2461 CATAACCAGG CTGAATCGCC CCATCATCCA GCCAGAAAGT GAGGGAGCCA CGGTTGATGA
 2521 GAGCTTTGTT GTAGGTGGAC CAGTTGGTGA TTTTGAACCT TTGCTTTGCC ACGGAACGGT
 2581 CTGCGTTGTC GGAAGATGC GTGATCTGAT CCTTCAACTC AGCAAAAAGTT CGATTATTTC
 2641 AACAAAGCCG CCGTCCCGTC AAGTCAGCGT AATGCTCTGC CAGTGTTACA ACCAATTAAC
 2701 CAATTCTGAT TAGAAAAACT CATCGAGCAT CAAATGAAAC TGCAATTTAT TCATATCAGG
 2761 ATTATCAATA CCATATTTTT GAAAAAGCCG TTTCTGTAAT GAAGGAGAAA ACTCACCGAG
 2821 GCAGTTCCAT AGGATGGCAA GATCCTGGTA TCGGTCTGCG ATTCCGACTC GTCCAACATC
 2881 AATACAACCT ATTAATTTCC CCTCGTCAAA AATAAGGTTA TCAAGTGAGA AATCACCATG
 2941 AGTGACGACT GAATCCGGTG AGAATGGCAA AAGCTTATGC ATTTCTTTCC AGACTTGTTT
 3001 AACAGGCCAG CCATTACGCT CGTCATCAAA ATCACTCGCA TCAACCAAAC CGTTATTTCAT
 3061 TCGTGATTGC GCCTGAGCGA GACGAAATAC GCGATCGCTG TTAAAAGGAC AATTACAAAC
 3121 AGGAATCGAA TGCAACCGGC GCAGGAACAC TGCCAGCGCA TCAACAATAT TTTCACCTGA
 3181 ATCAGGATAT TCTTCTAATA CCTGGAATGC TGTTTTCCCG GGGATCGCAG TGGTGAGTAA
 3241 CCATGCATCA TCAGGAGTAC GGATAAAATG CTTGATGGTC GGAAGAGGCA TAAATTCCGT
 3301 CAGCCAGTTT AGTCTGACCA TCTCATCTGT AACATCATTG GCAACGCTAC CTTTGCCATG
 3361 TTTCAGAAAC AACTCTGGCG CATCGGGCTT CCCATACAAT CGATAGATTG TCGCACCTGA
 3421 TTGCCCAGCA TTATCGCGAG CCCATTTATA CCCATATAAA TCAGCATCCA TGTTGGAATT
 3481 TAATCGCGGC CTCGAGCAAG ACGTTTCCCG TTGAATATGG CTCATAACAC CCCTTGTATT
 3541 ACTGTTTATG TAAGCAGACA GTTTTATTGT TCATGATGAT ATATTTTAT CTTGTGCAAT
 3601 GTAACATCAG AGATTTTGAG ACACAACGTG GCTTTGTTGA ATAAATCGAA CTTTGTCTGA
 3661 GTTGAAGGAT CAGATCACGC ATCTTCCCGA CAACGCAGAC CGTTCCGTGG CAAAGCAAAA
 3721 GTTCAAAATC ACCAACTGGT CCACCTACAA CAAAGCTCTC ATCAACCGTG GCTCCCTCAC
 3781 TTTCTGGCTG GATGATGGGG CGATTACGGC CTGGTATGAG TCAGCAACAC CTTCTTCACG
 3841 AGGCAGACCT CAGCGCTAGA TTATTGAAGC ATTTATCAGG GTTATTGTCT CATGAGCGGA
 3901 TACATATTTG AATGTATTTA GAAAAATAAA CAAATAGGGG TTCCGCGCAC ATTTCCCCGA
 3961 AAAGTGCCAC CTGA

Vektor mit Insert

ORIGIN

```

1 CGTAACTATA ACGGTCCTAA GGTAGCGAAA GCTCAGATCT CCCGATCCCC TATGGTGCAC
61 TCTCAGTACA ATCTGCTCTG ATGCCGCATA GTTAAGCCAG TATCTGCTCC CTGCTTGTGT
121 GTTGGAGGTC GCTGAGTAGT GCGCGAGCAA AATTAAAGCT ACAACAAGGC AAGGCTTGAC
181 CGAC AATTGAGTGGATTGCGGGGCACA GACGCCCAGGACCGCGCTTCCCACGTGGCGGAGGG
ACTGGGGACCCGGGCACCCGTCTGCCCTTCACCTTCCAGCTCCGCCTCCTCCGCGCGGACCCCGC
CCCGTCCCGACCCCTCCCGGGTCCCCGGCCAGCCCCCTCCGGGCCCTCCCAGCCCCCTCCCTTCTT
TTCCGCGGCCCCGCCCTCTCCTCGCGGCGCGAGTTTCAGGCAGCGCTGCGTCCTG CTGCGCACGTG
GGAAGCCCTCAATTCTAGACTCGAGC GGCCGCCACT
961 GTGCTGGATG ATCCGAGCTC
GGTACCCCGGTCATGGTGAGCAAGGGCGCCGAGCTGTTACCGGCATCGTGCCCATCCTGATCGAG
CTGAATGGCGATGTGAATGGCCACAAGTTCAGCGTGAGCGGCGAGGGCGAGGGCGATGCCACCTA
CGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGCAAGCTGCCTGTGCCCTGGCCCCACCCTGGT
GACCACCCTGAGCTACGGCGTGCAGTGCTTCTCACGCTACCCCGATCACATGAAGCAGCACGACTT
CTTCAAGAGCGCCATGCCTGAGGGCTACATCCAGGAGCGCACCATCTTCTTCGAGGATGACGGCAA
CTACAAGTCGCGCGCCGAGGTGAAGTTCGAGGGCGATAACCTGGTGAATCGCATCGAGCTGACCG
GCACCGATTTCAAGGAGGATGGCAACATCCTGGGCAATAAGATGGAGTACAAC TACAACGCCAC
AATGTGTACATCATGACCGACAAGGCCAAGAATGGCATCAAGGTGAACTTCAAGATCCGCCACAA
CATCGAGGATGGCAGCGTGCAGCTGGCCGACCACTACCAGCAGAATACCCCATCGGCGATGGCC
CTGTGCTGCTGCCGATAACCACTACCTGTCCACCCAGAGCGCCCTGTCCAAGGACCCCAACGAGA
AGCGCGATCACATGATCTACTTCGGCTTCGTGACCGCCGCGCCATCACCCACGG
AflII
CATGGATGAGCTGTACAAGTGAGCGGCCGCGACTCTAGATCATAATCAGCCATACCACATTTGCTT
AAGTAAGT GACTAGAACT TGTTTATTGC
1021 AGCTTATAAT GGTTACAAAT AAAGCAATAG CATCACAAAT TTCACAAATA AAGCATTTTT
1081 TTCACTGCAT TCTAGTTGTG GTTTGTCCAA ACTCATCTGC AGATCTGAAT TCATCTATGT
1141 CGGGTGCGGA GAAAGAGGTA ATGAAATGGC ATTATGGGTA TTATGGGTCT GCATTAATGA
1201 ATCGGCCAAC GCGCGGGGAG AGGCGGTTTG CGTATTGGGC GCTCTTCCGC TTCTCGCTC
1261 ACTGACTCGC TGCCTCGGT CGTTCGGCTG CGGCGAGCGG TATCAGCTCA CTCAAAGGCG
1321 GTAATACGGT TATCCACAGA ATCAGGGGAT AACGCAGGAA AGAACATGTG
AGCAAAAGGC
1381 CAGCAAAAGG CCAGGAACCG TAAAAAGGCC GCGTTGCTGG CGTTTTTCCA TAGGCTCCGC
1441 CCCCTGACG AGCATCACAA AAATCGACGC TCAAGTCAGA GGTGGCGAAA CCCGACAGGA
1501 CTATAAAGAT ACCAGGCGTT TCCCCCTGGA AGCTCCCTCG TGCCTCTCC TGTTCCGACC
1561 CTGCCGCTTA CCGGATACCT GTCCGCTTT CTCCCTTCGG GAAGCGTGCG GCTTTCTCAA
1621 TGCTCACGCT GTAGGTATCT CAGTTCGGTG TAGGTCGTTT GCTCCAAGCT GGGCTGTGTG
1681 CACGAACCCC CCGTTCAGCC CGACCGCTGC GCCTTATCCG GTAACATATCG TCTTGAGTCC

```

1741 AACCCGGTAA GACACGACTT ATCGCCACTG GCAGCAGCCA CTGGTAACAG GATTAGCAGA
 1801 GCGAGGTATG TAGGCGGTGC TACAGAGTTC TTGAAGTGGT GGCCTAACTA CGGCTACACT
 1861 AGAAGGACAG TATTTGGTAT CTGCGCTCTG CTGAAGCCAG TTACCTTCGG AAAAAGAGTT
 1921 GGTAAGCTCTT GATCCGGCAA ACAAACCACC GCTGGTAGCG GTGGTTTTTT TGTTTGCAAG
 1981 CAGCAGATTA CGCGCAGAAA AAAAGGATCT CAAGAAGATC CTTTGATCTT TTCTACGGGG
 2041 TCTGACGCTC AGTGGAACGA AAACCTACGT TAAGGGATTT TGGTCATGAG ATTATCAAAA
 2101 AGGATCTTCA CCTAGATCCT TTTGATCCTC CGGCGTTCAG CCTGTGCCAC AGCCGACAGG
 2161 ATGGTGACCA CCATTTGCCC CATATCACCG TCGGTACTGA TCCCGTCGTC AATAAACCGA
 2221 ACCGCTACAC CCTGAGCATC AAACCTCTTT ATCAGTTGGA TCATGTCGGC GGTGTCGCGG
 2281 CCAAGACGGT CGAGCTTCTT CACCAGAATG ACATCACCTT CCTCCACCTT CATCCTCAGC
 2341 AAATCCAGCC CTTCCCGATC TGTGAACTG CCGGATGCCT TGTCGGTAAA GATGCGGTTA
 2401 GCTTTTACCC CTGCATCTTT GAGCGCTGAG GTCTGCCTCG TGAAGAAGGT GTTGCTGACT
 2461 CATAACAGGC CTGAATCGCC CCATCATCCA GCCAGAAAGT GAGGGAGCCA CGGTTGATGA
 2521 GAGCTTTGTT GTAGGTGGAC CAGTTGGTGA TTTGAACTT TTGCTTTGCC ACGGAACGGT
 2581 CTGCGTTGTC GGGAAGATGC GTGATCTGAT CTTTCAACTC AGCAAAAGTT CGATTTATTC
 2641 AACAAAGCCG CCGTCCCGTC AAGTCAGCGT AATGCTCTGC CAGTGTTACA ACCAATTAAC
 2701 CAATTCTGAT TAGAAAAACT CATCGAGCAT CAAATGAAAC TGCAATTTAT TCATATCAGG
 2761 ATTATCAATA CCATATTTTT GAAAAAGCCG TTTCTGTAAT GAAGGAGAAA ACTCACCGAG
 2821 GCAGTTCCAT AGGATGGCAA GATCCTGGTA TCGGTCTGCG ATTCCGACTC GTCCAACATC
 2881 AATACAACCT ATTAATTTCC CCTCGTCAA AATAAGGTTA TCAAGTGAGA AATCACCATG
 2941 AGTGACGACT GAATCCGGTG AGAATGGCAA AAGCTTATGC ATTTCTTTCC AGACTTGTTT
 3001 AACAGGCCAG CCATTACGCT CGTCATCAA ATCACTCGCA TCAACCAAAC CGTTATTCAT
 3061 TCGTGATTGC GCCTGAGCGA GACGAAATAC GCGATCGCTG TTAAAAGGAC AATTACAAAC
 3121 AGGAATCGAA TGCAACCGGC GCAGGAACAC TGCCAGCGCA TCAACAATAT TTTCACCTGA
 3181 ATCAGGATAT TCTTCTAATA CCTGGAATGC TGTTTTCCCG GGGATCGCAG TGGTGAGTAA
 3241 CCATGCATCA TCAGGAGTAC GGATAAAATG CTTGATGGTC GGAAGAGGCA TAAATTCGCT
 3301 CAGCCAGTTT AGTCTGACCA TCTCATCTGT AACATCATTG GCAACGCTAC CTTTGCCATG
 3361 TTTCAGAAAC AACTCTGGCG CATCGGGCTT CCCATACAAT CGATAGATTG TCGCACCTGA
 3421 TTGCCCCACA TTATCGCGAG CCCATTTATA CCCATATAAA TCAGCATCCA TGTTGGAATT
 3481 TAATCGCGGC CTCGAGCAAG ACGTTTCCCG TTGAATATGG CTCATAACAC CCCTTGATT
 3541 ACTGTTTATG TAAGCAGACA GTTTTATTGT TCATGATGAT ATATTTTAT CTTGTGCAAT
 3601 GTAACATCAG AGATTTTGAG ACACAACGTG GCTTTGTTGA ATAAATCGAA CTTTGTGCTGA
 3661 GTTGAAGGAT CAGATCACGC ATCTCCCGA CAACGCAGAC CGTTCCGTGG CAAAGCAAAA
 3721 GTTCAAAATC ACCAACTGGT CCACCTACAA CAAAGCTCTC ATCAACCGTG GCTCCCTCAC
 3781 TTTCTGGCTG GATGATGGGG CGATTACGGC CTGGTATGAG TCAGCAACAC CTTCTTCACG
 3841 AGGCAGACCT CAGCGCTAGA TTATTGAAGC ATTTATCAGG GTTATTGTCT CATGAGCGGA
 3901 TACATATTTG AATGTATTTA GAAAAATAAA CAAATAGGGG TTCCGCGCAC ATTTCCCCGA
 3961 AAAGTGCCAC CTGA

14. Curriculum vitae



Persönliche Information

Name	Eva Maria Hoffmann, Mag. rer. nat.
Adresse	Hauptstrasse 154, 8401 Kalsdorf
Telefonnummer	0699 11051231
Email	eva.hoffmann@medunigraz.at
Geburtsdatum	22. März 1981
Familienstand	ledig
Nationalität	Österreicherin

Wissenschaftliche Ausbildung

Seit 05/ 2012	Autorin beim »Schwarzkopf & Schwarzkopf Verlag« (Berlin)
10/ 2009 – 10/ 2012	PhD-Programm “Molecular Medicine”, Medizinische Universität Graz, Institut für Humangenetik. “Establishment of new methods for isolation and detection of circulating tumor cells and cell-free circulating tumor DNA in cancer patients”. Betreuer: Univ.-Prof. Michael Speicher, Dr. med. Assoz. Prof. Priv.-Doz. Jochen Geigl, Dr.med.
12/ 2007	Graduierung zur Mag^{ra}.rer.nat.
11/ 2006 – 12/ 2007	Magisterstudium "Molekulare Mikrobiologie", Karl-Franzens-Universität Graz. Hauptfächer: Molekularbiologie, Genetik, Molekulare Biotechnologie

Diplomarbeit: “Development of a new expression system in *Pichia pastoris*”. Abschluss mit ausgezeichnetem Erfolg am Institut für Molekulare Biotechnologie, Technische Universität, Graz.

Betreuer: Univ.-Prof. Helmut Schwab, Dipl.-Ing. Dr.tech.

Ass. Prof. Harald Pichler, Dipl.-Ing. Dr.tech.

10/ 2001 – 11/ 2006

Bakkalaureats-Studium

"Molekularbiologie an der Karl-Franzens-Universität Graz

Hauptfächer: Molekularbiologie, Zoologie, Botanik

Bakkalaureats-Arbeiten: "Meningitis", "Genetherapy in cancer"

10/ 1999 - 03/ 2001

Medizinstudium, Medizinische Universität Graz.

1995 - 1999

BG/ BRG Klusemannstraße, Graz, Austria

Hauptfächer: Medien und Kommunikation

Abschluss mit ausgezeichnetem Erfolg

Berufliche Praxis

04/ 2009 – 10/ 2012

PhD-Studentin am Institut für Humangenetik,
Medizinische Universität Graz.

PhD-Program: “Molecular Medicine”

2009 – 2010

Mitarbeit an einem Grazer Studentenmagazin (Schreiben von
Kolumnen)

10/ 2008 - 10/ 2009

Doctoral school: "Molecular Medicine and Inflammation",
Medizinische Universität Graz

07/ 2006 - 09/ 2006

Praktikum am Institut für Molekulare Biotechnologie, Graz,
“Plasmidtransfer via Triparental Mating”

2004 – 2007

Promotionstätigkeit (Flyer-Verteilaktionen, Event-Mitarbeit, Verkauf)
für Hutchison 3G

Sonstige Kenntnisse

Sprachen

Deutsch (Muttersprache), English (sehr gut)

Computer-Kenntnisse

MS-Office (Word, Excel, PowerPoint)

Hobbies

Lesen	vor allem Bücher österreichischer Schriftsteller (Thomas Glavinic, Daniel Glattauer, Linda Stift,...)
Schreiben	Veröffentlichungen: »Anfangen, Abgebrochen, Neu erfunden«; (April 2013, im Schwarzkopf-Verlag) »Fräulein Monas Suche nach dem Sinn« , (Oktober 2013, im „Der Kleine Buchverlag“)
Kochen	vor allem Indisch