Aus der II. Medizinischen Klinik
der Medizinischen Fakultät Mannheim
(Direktor: Prof. Dr. med. Matthias Ebert)

Multiplex-PCR and deep sequencing for mutation detection in circulating cell-free DNA of colorectal cancer patients

Inauguraldissertation
zur Erlangung des medizinischen Doktorgrades
der
Medizinischen Fakultät Mannheim
der Ruprecht-Karls-Universität
zu
Heidelberg

vorgelegt von
Simon Herrmann
aus
Tübingen
2019

Dekan: Prof. Dr. med. Sergij Goerdt

Referent: Prof. Dr. med. Matthias Ebert

TABLE OF CONTENTS

ABBREVIATIONS

5-FU	5-fluorouracil
AFP	alpha feto-protein
ALAT	alanine aminotransferase
AP	alkaline phosphatase
ARMS-PCR	amplification refractory mutation system PCR
ASAT	aspartate aminotransferase
BCLC	Barcelona clinic liver cancer
BEAMing	beads, emulsions, amplification and magnetics
BMP	bone morphogenetic protein
bp	base pairs
CA19-9	cancer antigen 19-9
CAPOX	capecitabine + oxaliplatin
CCC	cholangiocellular carcinoma
CCC	cholangiocellular carcinoma
CEA	carcinoembryonic antigen
CIN	chromosomal instability
cirDNA	circulating cell-free DNA
cirNA	circulating cell-free nucleic acids
cirRNA	circulating cell-free RNA
COSMIC	Catalogue of Somatic Mutations in Cancer
CRC	colorectal cancer
CRP	C-reactive protein
CTC	circulating tumor cells
ctDNA	circulating cell-free tumor DNA
ddPCR	digital droplet PCR
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
dsDNA	double-stranded DNA
EC	esophageal cancer
EDTA	ethylenediaminetetraacetic acid
EGF	epidermal growth factor
EGFR	epidermal growth factor receptor
EMT	epithelial-to-mesenchymal transition
ERK	extracellular-signal-regulated kinases
EUSTM	European Society for Translational Medicine
FAP	familial adenomatous polyposis
FBS	fetal bovine serum
FFPE	formalin-fixed paraffin embedded
FOLFIRI	5-fluorouracil + leucovorin + irinotecan
FOLFIRINOX	5-fluorouracil + leucovorin + oxaliplatin + irinotecan
FOLFOX	5-fluorouracil + oxaliplatin
FW	forward
GC	gastric cancer
GGT	gamma-glutamyl transferase
GIST	gastrointestinal stroma tumor
GI-tumors	gastrointestinal tumors
HCC	hepatocellular carcinoma
HNPCC	hereditary non-polyposis colorectal cancer
INR	international normalized ratio
LDH	lactate dehydrogenase
LV	leucovorin
MAC	mutant allele concentration
MAF	mutant allele frequency
MALIBU	Mannheim Liquid Biopsy Unit
MAPK	mitogen-activated protein kinases
mCRC	metastasized colorectal cancer
MMR	mismatch repair
MRD	minimal residual disease
MSI-(H) or(L)	microsatellite instability (high) or (low)
MSS	microsatellite stability

NCBI	National Center for Biotechnology Information
NCI	National Cancer Institute
ND	not done
NET	neuroendocrine tumors
NGS	next-generation sequencing
NIPT	non-invasive prenatal testing
NSCLC	non-small cell lung cancer
PBS	phosphate buffered saline
PC	pancreatic cancer
PCC	pearson correlation coefficient
PCR	polymerase chain reaction
PFS	progression-free survival
PI3K	phosphoinositide 3-kinase
PIP2	phosphatidylinositol-4,5-biphosphate
PIP3	phosphatidylinositol-3,4,5-trisphosphate
qPCR	quantitative real-time PCR
RBC	red blood cell
RECIST	Response evaluation criteria in solid tumors
REV	reverse
RNA	ribonucleic acid
SCC	spearman correlation coefficient
SIC	small intestinal cancer
TAE	tris base, acetic acid and EDTA
TEP	tumor-educated blood platelets
TGF	transforming growth factor
TGFβ	Transforming growth factor beta
TLR9	toll-like receptor 9
TM	translational medicine
TMB	tumor mutational burden
TNM	tumor, node, metastasis
TS	thymidylate synthase
UAS	universal adapter sequence
UICC	Union for International Cancer Control
UMI	Unique molecular identifier
VEGF	vascular endothelial growth factor
VEGFR	VEGF receptor
WBC	white blood cells
WES	whole-exome sequencing
WGS	whole-genome sequencing
WHO	World Health Organization
WT	wild type

Gene abbreviations according to the HUGO Gene Nomenclature Committee (HGNC) were used.

PROLOGUE

Parts of this thesis were previously published in Molecular Oncology in the article “Detection of mutational patterns in cell-free DNA of colorectal cancer by custom amplicon sequencing” in June 2019.

1 INTRODUCTION

1.1 Gastrointestinal Cancer

Cancer is defined by the WHO as a disease *"characterized by the growth of abnormal cells beyond their usual boundaries that can then invade adjoining parts of the body and/or spread to other organs."*[1]

Despite remarkable improvements in diagnostics and therapy throughout the last decades, malignant neoplasms are still the second most common cause of death in Germany and other developed countries, only exceeded by cardiovascular diseases.

Taking a closer look at the most frequent tumor origins, gastrointestinal neoplasms stand out, representing 5 of the 10 cancer types with the highest mortality (colon, rectum, pancreas, stomach, liver).[2] Besides hepatocellular carcinoma (HCC) and squamous cell carcinoma of the esophagus and anus the overwhelming majority of gastrointestinal malignancies are adenocarcinomas deriving from their respective organ of origin.[3]

1.1.1 Colorectal Cancer

Colorectal Cancer (CRC) is a leading cause of cancer morbidity and mortality worldwide with now more than 1.8 million patients diagnosed every year and almost every second dying of this disease.[4]

In Germany CRC has the second highest incidence in females, whereas in males it now shares second place with lung cancer.[5] The incidence rates are multiple times higher in the developed countries of Europe, North America and Asia compared to third world countries. These great differences can be partially explained by the demographic structure in these countries with its aging population and life-style risk factors, such as smoking, obesity, physical inactivity, alcohol consumption, a nutrition rich in high calorie food and red meat.[6] An additional risk factor for CRC is inheritance, which may be responsible for up to 30% of all cases.[7] There are multiple known monogenetic syndromes associated with CRC, most notably the Lynch syndrome. Here, mutations in mismatch repair genes (MMR) lead to microsatellite instability and consequently to cancer not only of the intestines, but of several extra-intestinal organs.[8]

Although mortality rates have been falling over the last decades in Germany, the United States and Western Europe, owing to improved treatment, increased awareness and early detection, the number of new diagnoses were still rising.[9]

As most CRCs arise from adenomatous polyps and the development from adenoma to an invasive form requires more than ten years, it can be prevented by early diagnosis and endoscopic removal.[10] After colonoscopy as part of screening programs for precancerous lesions has been introduced in the last decade, a drop in incidence rates most notably in distal colon cancer in adults aged 50 years and older is now observed[11,12].

Progress in surgical procedures, new treatment regimens, such as neoadjuvant protocols and advanced combination chemotherapy have contributed to higher survival rates and prolonged overall survival.[11] In metastasized CRC (mCRC) oxaliplatin and irinotecan improved classic 5-fluorouracil-based (5-FU) protocols[13]. Targeted therapy as a novel statue of therapy especially in patients with advanced and inoperable disease improved their overall outcome. Monoclonal antibodies targeting the epidermal-growth factor receptor (EGFR) cetuximab and panitumumab or the vascular endothelial growth factor (VEGF) and the VEGF-receptor (VEGFR) bevacizumab and aflibercept are now regularly applied.[14,15] Whether anti-EGFR therapy is eligible depends on the mutational status of RAS genes and BRAF.[16–18] Especially patients with left-sided CRC benefit from this treatment.[19]

1.2 Translational medicine, personalized medicine and their promises

The European Society for Translational Medicine (EUSTM) defines translational medicine (TM) as *"an interdisciplinary branch of the biomedical field supported by three main pillars*: *benchside*, *bedside and community. The goal of TM is to combine disciplines, resources*, *expertise*, *and techniques within these pillars to promote enhancements in prevention*, *diagnosis*, *and therapies*."[20]

The National Cancer Institute (NCI) describes precision or personalized medicine as *"a form of medicine that uses information about a person's genes, proteins, and environment to prevent, diagnose, and treat disease."*[21]

In the context of cancer, this means collecting information about the patient's tumor, that may influence his treatment options.

1.3 Carcinogenesis in the large intestine

More than 60 years ago it has been postulated that colorectal cancer develops out of an adenoma, a benign polyp of the intestinal mucosa.[22] The adenoma-carcinoma sequence described by Vogelstein and Fearon in 1990 exemplifies this carcinogenic process.[23] Over time the adenoma acquires mutations in genes regulating cell cycle, growth and DNA repair and thus progresses gradually to a dysplastic adenoma and finally an invasive adenocarcinoma. Activation of proto-oncogenes or silencing of tumor-suppressor genes provide the altered cells a growth advantage.[24] At first, mutations in the APC gene and loss of the second allele initiate the transformation of the mucosa (Figure 1-1). Thereafter KRAS is activated through mutation. Further progression needs mutations in or loss of genes such as DCC, PIK3CA, SMAD4 and TP53. Often, even larger deletions of the respective chromosomal regions (e.g. 5q, 18q, 17p) occur. Another defining characteristic is the aneuploidy of those tumor cells. This can be summarized as the chromosomal instability (CIN) pathway and accounts for approximately 80% of all sporadic CRCs.[3,25]

Genomic instability plays a key role in carcinogenesis, as it allows rapid accumulation of genetic changes. The second major model for carcinogenesis in CRC accounts for genomic instability due to DNA mismatch-repair deficiency, the microsatellite instability pathway (MSI) (discussed in detail below).[26]

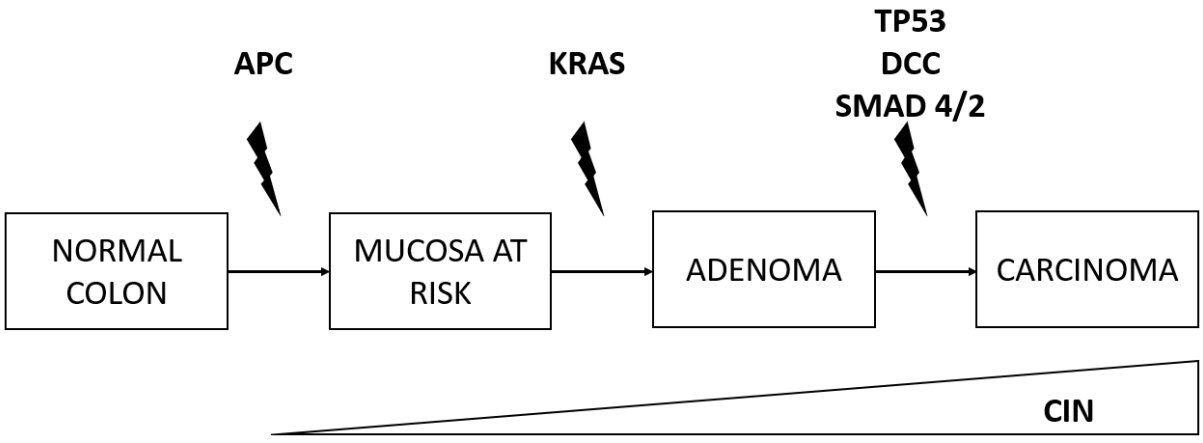

Figure 1-1. The adenoma-carcinoma sequence in the large intestine. The development to a malignant tumor requires multiple “hits”, i.e. mutations, that transform cells to tumor cells. This is in part based on the model of tumorigenesis by Knudson.[27]

1.4 The mutational spectrum of colorectal cancer

Colorectal cancer is a tumor that bears an intermediate to high tumor mutational burden (TMB) and often has a characteristic mutational spectrum.[28] In the following, relevant pathways implicated in the carcinogenesis of CRC will be described.

1.4.1 APC and Wnt signaling

Wnt signaling is one of the dominant pathways involved in colorectal carcinogenesis as it is altered in more than 90% of all CRCs. Adenomatous polyposis coli (APC), a large tumor-suppressor gene on Chromosome 5q, is an essential component of the β-catenin destruction complex and its mutation rate ranges between 60-80% in large scale sequencing studies.[29–31] β-catenin is the effector of the canonical Wnt pathway and functions as a transcriptional co-activator and, if not deactivated, causes a continuous proliferation of the cell.[32] Its coding gene CTNNB1 is mutated in 6% of cases and this occurs mutual exclusive with the APC mutation.[33]

The involvement of APC in CRC became evident through the heterozygous germline mutation found in familial adenomatous polyposis (FAP). FAP is an obligate precancerosis of CR with patients developing hundreds of polyps and almost always cancer in their second decade of life.[34] In sporadic CRC mostly truncating mutations of APC occur combined with a loss of the second allele. The majority of mutations is distributed over a few hotspots (Figure 1-2) (cancer.Sanger.ac.uk).[35]

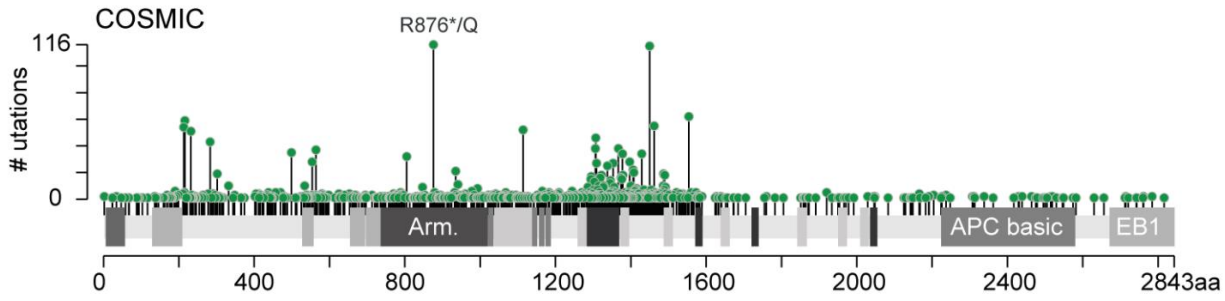

Figure 1-2. APC mutational hotspots in colorectal cancer. [35]

1.4.2 P53 – guardian of the genome

TP53 is a tumor suppressor gene encoding the transcription factor, p53, and the most commonly mutated gene in human cancer. p53 plays a crucial role in cell cycle regulation and DNA damage repair. It thwarts neoplastic development through three

intermeshing mechanisms quiescence, senescence and apoptosis. p53 monitors stress levels and, if needed, directs towards one of those three pathways. This stress can be DNA damage, anoxia or high oncoprotein activity. When cells go into quiescence the cell cycle is paused in the G1 phase giving time to repair possible errors. The cell cycle can even be stopped permanently, which accounts for the senescence mechanism. In addition, p53 stimulates multiple genes belonging to different DNA repair mechanisms, if irreversible damage to the cell has occurred apoptosis is induced by p53.[36] Deficient p53 causes DNA damage to remain unrepaired and cells divide and further accumulate mutations leading to malignant transformation.[36]

In CRC TP53 is modified in 63%.[33] Mutant TP53 is typically accompanied by copy number variations and aneuploidy. The mutational spectrum covers the whole gene with a few distinct hotspots standing out.[33,36]

Inherited TP53 mutations cause the Li-Fraumeni syndrome. Patients are at high risk to experience multiple neoplasms at young age, including breast cancer, sarcomas, adrenal cortical carcinoma and leukemia and to a lesser extent gastrointestinal neoplasms.[37,38]

1.4.3 Microsatellite instability (MSI) – the mutator pathway

MSI can either be hereditary due to germline mutations or sporadic due to somatic changes in mismatch repair genes.[26,39] MMR is the DNA damage repair mechanism, which plays a dominant role in colorectal carcinogenesis. Relevant genes are MLH1, MSH2, MS6, MLH3 and PMS2. Together they recognize and correct small insertions, deletions and misincorporated bases. If both alleles of an aforenamed gene are affected, errors in replication will no longer be fixed, resulting in high rates of point mutations throughout the genome,[40] especially in microsatellites,[41] which are non-coding short tandem repeats. Their pattern is unique and differs in between individuals.[42] Subsequently mutations in regulator genes such as TGFβRII, BAX, TCF-4 and others accumulate and contribute to the carcinogenic process. Those genes have microsatellite sequences within their encoding regions and are thus susceptible to acquire alterations in this context.[43] MSI-tumors are typically euploid, bear less chromosomal aberrations and show lower frequencies of KRAS, TP53 and APC mutations, whereas BRAF V600E is associated with sporadic MSI-tumors.[44] According to results of histological testing for microsatellite instability, high, low and no MSI can

be distinguished (MSI-H, MSI-L and MSS). Efforts to determine the MSI-status with the use of DNA-sequencing are under investigation and may replace the histological markers in near future.[45]

Epigenetic silencing is responsible for most sporadic forms. The promoter regions of MMR genes, predominantly MLH1, are hypermethylated and thus will be less transcribed.[39] Lynch syndrome, the autosomal dominant hereditary non-polyposis colorectal cancer (HNPCC), constitutes about 2-4% of all CRC cases. The underlying genetic defect is a heterozygous germline mutation in one of the MMR genes, leading to high susceptibility to CRC, endometrial cancer, stomach cancer and others. The lifetime prevalence of CRC is up to 80% in affected persons.[8]

MSI-status is not only relevant in Lynch syndrome. MSI-tumors are less prone to lymphatic and hematogenous metastasis.[46] A systematic review confirmed a prognostic advantage of the MSI-phenotype.[47] This is especially the case for stage II MSI-tumors and whether those patients benefit from an adjuvant treatment with 5-FU remains questionable.[48,49] In the metastatic situation the MSI-phenotype is uncommon, with rates as low as 4%. Here, the impact on prognosis is unclear and it might even be negative.[44,50]

1.4.4 MAPK signaling cascade

The mitogen-activated-protein-kinases (MAPK) pathway is of great importance for cell proliferation. There are three major tracks with the extracellular-signal-regulated kinases (ERK-MAPK) pathway being crucial in CRC (Figure 1-3). Downstream of receptor tyrosine kinases such as EGFR, the ERK-MAPK pathway mediates cell growth, survival and invasion. Upon binding of EGF to its receptor, the EGFR-RAS-RAF-MEK-ERK axis is activated. ERK is translocated into the nucleus and phosphorylates various transcription factors, e.g. c-Jun or c-Fos.[51]

Genes of this pathway are frequently mutated or overexpressed in CRC. As they function as proto-oncogenes, they commonly acquire activating mutations in distinct hotspots. KRAS, NRAS and BRAF are altered in 41%, 6% and 13% of cases. Mutations in these genes are mutual exclusive. Hence, alterations in this pathway can be found in at least every second patient.[33]

BRAF mutations are of upcoming interest, as they have multiple implications in mCRC. First, BRAF V600E renders affected cells insensitive to anti-EGFR-therapy.[52] Secondly, it is associated with a very poor prognosis, right-sided colon cancer and

peritoneal carcinosis and on the molecular level with the MSI-phenotype.[44,53,54] BRAF V600E accounts for 80% of all alterations. In contrast, non-V600E mutations are associated with longer overall survival, even when compared to BRAF WT.[55]

1.4.5 PI3K-AKT signaling

Tightly linked with MAPK signaling, the PI3K-AKT pathway also includes numerous cancer genes (Figure 1-3). After growth factors bind to their receptors, among others EGFR, phosphoinositide 3-kinases (PI3K) are activated, which convert phosphatidylinositol-4,5-biphosphate (PIP2) to phosphatidylinositol-3,4,5-trisphosphate (PIP3). PTEN can inhibit PI3K. Upon conversion AKT is recruited to the plasma membrane and phosphorylated by phosphoinositide-dependent kinase 1 (PDK1) and PKD2. AKT has multiple downstream substrates, notably the mTOR complex. The effectors regulate cell growth, proliferation, differentiation and apoptosis.[56]

Except for PTEN all genes are proto-oncogenes. PIK3CA encodes the p110α catalytic subunit of class IA PI3Ks.[56] In CRC PIK3CA is activated by mutations in more than 20% of patients, stating its relevance in this pathway and the oncogenic impact. PTEN is affected in 8% of cases, either the tumor suppressor gene is homo-deleted or mutations in hotspots are identifiable. Germline defects in PTEN are associated with the rare Cowden syndrome, a hamartoma polyposis syndrome, which also bears the risk of developing cancer of the breast, thyroid and skin.[57] Alterations in this genes co-occur with each other and KRAS, NRAS and BRAF.[33]

Though mechanistically activating alterations in PIK3CA could confer resistance to anti-EGFR treatment, past results yielded controversial results. For CRC with exon 20 mutations lower response rates have been described and missing mutations correlated with slightly better clinical response.[18,58] In conclusion, current data suggests only a marginal predictive role for PIK3Ca mutations.[59]

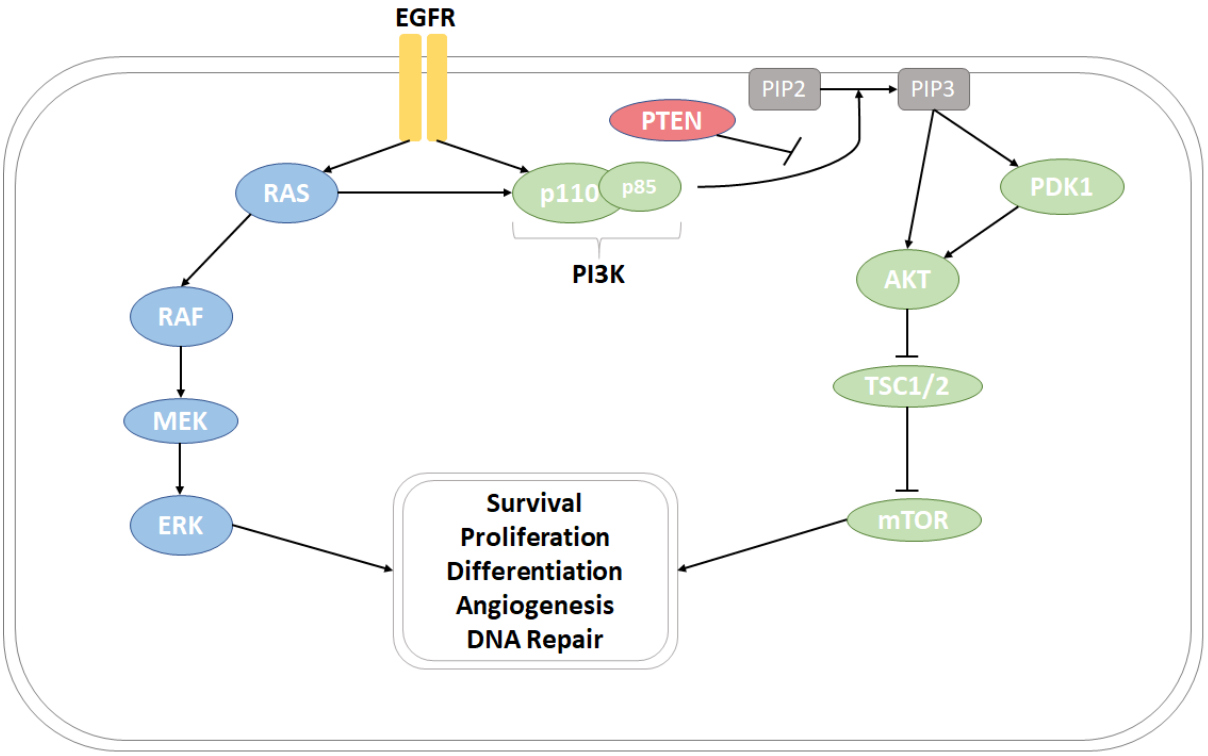

Figure 1-3. Simplified signaling cascades of the ERK-MAPK-pathway and the PI3K-AKT-pathway, their interrelations and their effects in cancer. After activation of upstream receptors effector proteins are phosphorylated and eventually nuclear transcription factors will be regulated in their activity. In cancer both pathways play an eminent role in carcinogenesis and tumor progression. PTEN is the only tumor suppressor of both pathways, the other proteins are often uncoupled from upstream signals by activating mutations.

1.4.6 TGFβ signaling

Transforming growth factor beta (TGFβ) signaling influences proliferation and differentiation in multiple cell types. It has been termed “the molecular Jekyll and Hyde of cancer”, as its effects can be quite different, either tumor suppression or oncogenesis.[60]

In CRC this pathway is commonly impaired and its growth inhibiting effect is lost, which contributes to carcinogenesis.[25,60] In many late-stage tumors its impact is shifted and epithelial-to-mesenchymal transition (EMT) is induced, thus promoting migration, invasion and dissemination.[61] Oncogenic effects are partly mediated through interaction with the ERK-MAPK, PI3K/AKT and non-canonical WNT pathways.[62]

After binding of TGFβ to TGFBR2, a group of proteins called SMADs are activated. These central mediators form complexes and translocate to the nucleus, where they convey their effects by activating transcription factors or they act as transcription factors by themselves.[60]

Familial juvenile polyposis, another hamartoma syndrome which leads to high susceptibility to polyps and CRC, is caused by germline defects of SMAD4.[63] This is also the most frequently altered gene of this pathway in CRC (16%). It can either be deleted or carry hotspot mutations. The loss of SMAD4 has been shown to cause resistance to 5-FU based therapy regimes through activation of the PI3K/AKT pathway.[64] Other genes affected are SMAD2, SMAD3 and TGFBR2 (6%, 5% and 5%).[33] TGFBR2 was previously believed to be defect in up to 30% of all cases[62], yet recent large sequencing studies gained conflicting numbers with rates down to 2-4%.

1.4.7 Other altered genes in CRC

FBXW7

FBXW7 encodes a component of the SCF E3 ubiquitin ligase. It is an important tumor suppressor as it is responsible for the degradation of many oncogenic proteins. The third and fourth WD40 domain of FBXW7 include a highly conserved arginine-containing region (R465, R479, R505), which is essential for binding to its substrates.[65] Almost half of the mutations in FBXW7 are detected there. 14% of CRCs carry mutations in FBXW7.[33]

GNAS

As a proto-oncogene GNAS can either be amplified or activated by distinct mutations. The latter is found in a large proportion of villous adenoma and about 4% of CRCs.[33,66] GNAS encodes for an α-subunit of a G-protein. This transduces signals from various hormone and growth factor receptors and influences diverse signaling pathways. In the large intestine its activation promotes proliferation to some extent through Wnt and MAPK signaling.[67]

1.4.8 Drug metabolism of 5-FU

Presently, it is not possible to predict whether a patient will benefit from a chemotherapy or if he only will endure adverse effects and when he will have a recurrence or progress of his disease. For targeted therapeutics the analysis of respective pathway components can provide valuable information about the tumor's therapy response.

Almost every patient with a CRC receives 5-FU and even those who initially respond will eventually develop a progress.[68] 5-FU is an antimetabolite, which damages all

dividing cells. 5-FU is an uracil-analogue with a fluorine atom at the C-5 position. After intracellular activation it inhibits the thymidylate synthase (TS), compromising DNA synthesis and repair or it is incorporated into DNA and RNA, thereby causing strand breaks.[69]

A reasonable attempt to find out about resistance mechanisms is to look at the corresponding genes relevant for the metabolism of 5-FU. In CRC these are infrequently altered and do not show hotspots. However, most sequenced tumor samples so far are pretherapeutic, thus the mutational spectrum under therapy may differ and tumors may acquire a resistance mechanism through mutation of genes responsible for drug metabolism.

Multiple polymorphisms in the 5-FU metabolism have been described and associated with varying responses to 5-FU treatment.[70,71] Past efforts to integrate biomarkers for 5-FU response into routine management have failed. Low intratumoral mRNA expression of TS, thymidylate phosphorylase and dihydropyrimidine dehydrogenase were associated with longer overall survival in CRC patients. Copy numbers changes of the TS gene (TYMS) conferred resistance to 5-FU,[72,73] however several following studies reported ambiguous results.[74]

In conclusion, signaling pathways build a complex network to influence the cells' fate. In cancer deregulation of those is the basis of the carcinogenic process and tumor cells acquire multiple genetic changes affecting them. Hence, studying mutational patterns under therapy offers a chance to understand their roles in cancer and possible resistance mechanisms.

1.5 "Liquid biopsy" – circulating nucleic acids in cancer diagnostics

Through sampling of blood plasma or other bodily fluids tumor material (e.g. cells or nucleic acids) can be acquired in a minimally invasive manner and be further analyzed. These procedures are summarized under the term "liquid biopsy".

Different substrates have been found to obtain relevant information about the patient's tumor, including circulating tumor cells (CTCs), tumor-educated blood platelets (TEPs) and cell-free RNA & DNA.

For this work the focus will be the relevance of circulating cell-free DNA (cirDNA) in the blood plasma of patients with gastrointestinal cancers, in particular with CRC.

1.5.1 Circulating cell-free DNA in blood plasma

It has been known for a long time that free-floating DNA fragments can be found in plasma samples. First described in 1948 by Mandel and Métais[75], progress in the field of liquid biopsies was slowed down due to a lack of suitable methods for further analyses.

In 1977 a first report of elevated cirDNA-levels in cancer patients was published.[76] However, elevated levels can also be observed in patients with acute trauma, stroke, infections or after exercise.[77] In 1989, Stroun and colleagues[78] identified cirDNA-molecules deriving from the tumor itself and soon after cell-free tumor DNA (ctDNA) was detected in other bodily fluids such as urine, cerebrospinal fluid, pleural fluid or stool samples.[77]

Origin of cirDNA

cirDNA is released through apoptosis and necrosis of cells and partly by active secretion.[79–81] Among others exosomes, a form of extracellular vesicles which are large nucleic acid-lipoprotein complexes, are actively secreted (Figure 1-4).[80] They both carry RNA and DNA molecules. Cellular sources of cirDNA are predominantly white blood cells (WBCs) and in cancer patients WBCs, malignant cells and cells of the tumor microenvironment.[82,83] In the blood stream cirDNA is able to form complexes with albumin and other serum proteins or attach to cell surfaces of circulating cells.[84,85]

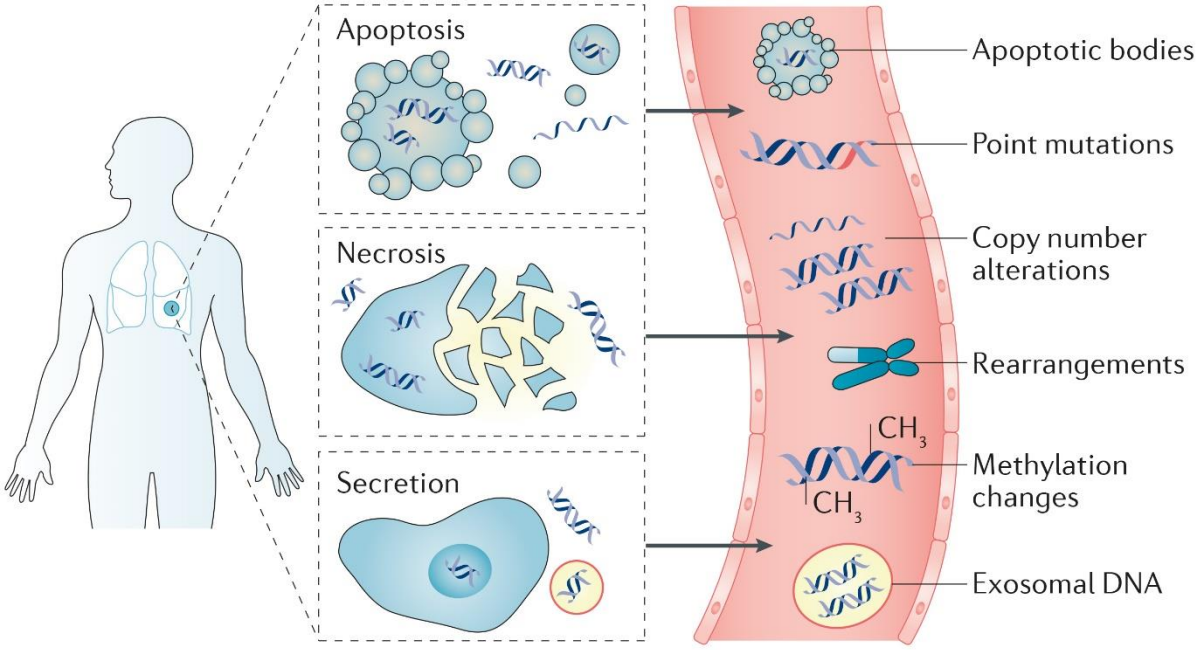

Figure 1-4. Origins of cirDNA and possible alterations detectable , as outlined by Wan et al.[77]

Size distribution in blood plasma

In healthy individuals cirDNA fragments are predominantly 166bp long, constituting the length of a chromatosome unit. A nucleosome consists of a histone octamer with DNA of ~147bp length wrapped around and forms together with a linkage DNA a chromatosome.[82] Chromatosome units are products of apoptotic DNA degradation, which is why apoptosis is thought to be the main source of cirDNA.[79–81] Fragment sizes show a 10bp ladder pattern due to nuclease cleaving at predisposed sites.[81,82] Also, large fragments (>1000bp) can be detected, typically deriving from necrotic cell death or exosomes.[79,80]

In cancer the situation is different. While the peak length is still ~166bp, studies have shown that cirDNA in tumor patients has a higher fragmentation (<145bp) (Figure 1-5).[86,87] The underlying reason for this shorter fragment length remains unclear. Extensive cleavage by nucleases is discussed as well as a different origin of those fragments or other release mechanisms.[88]

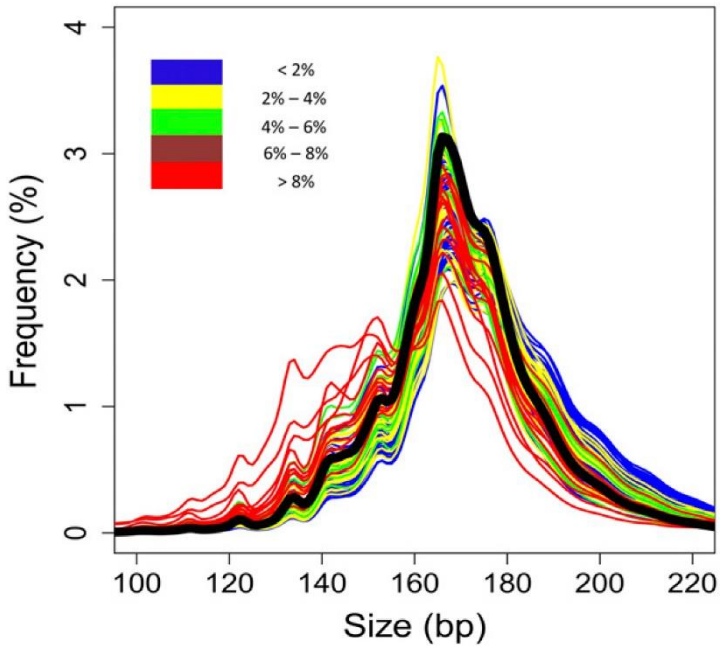

Figure 1-5. Length distribution of cirDNA in HCC patients and healthy individuals , according to Jiang et al.[86] The median size distribution of 32 healthy individuals is drawn as a broad black line. With higher ctDNA fractions the median size is shifted to smaller fragments (red lines).

Functional role of cirDNA

cirDNA molecules are recognized by Toll-like receptor 9 (TLR9), a membrane-bound protein which is commonly expressed by immune cells, implying potential immunomodulatory effects.[89] The fact that, for example in systemic lupus erythematosus antibodies against double-stranded DNA (anti-dsDNA) play a key role

in the pathogenesis supports this hypothesis.[90] Moreover, cirDNA can also be internalized by cells and is capable of transfection, i.e. the horizontal transfer of genes into other cell's genomes. In cancer this concept is termed genometastasis.[91]

cirDNA- and ctDNA-levels

Between different cancer types levels of cirDNA and ctDNA vary greatly and higher levels are observed in advanced stages.[92,93] The mutant allele fraction (MAF), i.e. the fraction of ctDNA within the total amount of cirDNA, ranges between 0% and up to 95%, highlighting the interindividual heterogeneity. This also applies to the concentration of ctDNA.[92–95] Furthermore, differing levels of ctDNA in serial samples demonstrate, that intraindividual heterogeneity exists as well.[96,97]

For how long cirDNA circulates in the blood stream is not well determined. Current observational data for the half-life time of cirDNA molecules varies between 16 minutes and almost 2.5 hours.[98,99] Underlying mechanisms include degradation by circulating nucleases, phagocytosis, but also hepatic and renal clearance.[77]

As aforementioned, the current clinical condition (e.g. infection, physical stress or a malignant disease) affects levels of cirDNA.[77] In pathologic conditions, the binding to serum proteins and cell surfaces can also be altered and the elimination of cirDNA impaired.[77,84,85] The composition of cirDNA can also fluctuate, as the DNA molecules can either circulate as chromatosome units, bind to proteins, cells and other macromolecules or be internalized into microvesicles. With current methods it is not possible to accurately determine the respective proportions.[88]

In cancer the tumor entity and the overall tumor burden accounts for the high variability in cirDNA- and ctDNA-levels.[92,100] Its specific location, tumor vascularization and cell death rate are thought to add to this variability.[92,101] The time point of blood sampling is another critical aspect. As cancer therapeutics optimally hamper tumor growth and cause extensive tumor cell death, higher levels of cirDNA and ctDNA can be observed days after treatment application, but also the absence of such spikes has been reported.[102,103] Whether cirDNA-levels follow a clear circadian rhythm is unknown, but variations are noticed.[104]

When comparing data of different groups, analytical differences such as sample preparation, extraction and detection methods need to be considered (further discussed below). All these factors explain the strong inter- as well as intraindividual heterogeneity in cirDNA-levels.

1.5.2 Methodological aspects of cirDNA analysis

Sample preparation

Plasma is preferred over serum as matrix of choice. Higher levels of cirDNA are found in serum samples and this is due to extensive lysis of WBCs.[105]

To ensure proper polymerase chain reaction (PCR) amplification, EDTA is the anticoagulant of choice, as heparin might inhibit PCR reactions.[106] Besides, EDTA proved to stabilize cirDNA and blood cells for up to one day without causing significant changes of total cirDNA-levels.[107] If plasma isolation within 6-24 hours cannot be guaranteed, the use of special storage tubes should be considered.[107] Centrifugation protocols for plasma processing unanimously favor a two-step approach, which minimizes contamination with genomic DNA.[108] Isolated plasma can be stored at -20°C or -80°C and repeated freeze-thaw cycles should be avoided, thus sample aliquotation is preferred.[108]

cirDNA isolation

Multiple commercial kits for DNA extraction in blood samples are available. Initially the QIAamp DNA Blood Mini kit (Qiagen) was shown to have the best recovery,[109] however, now specially designed purification kits for cirDNA are available and those perform better in the case of very low-level concentrations.[110,111] An important side effect of cirDNA extraction is that different methods or kits preferentially recover DNA fragments of different sizes.[110]

Accurate determination of cirDNA-levels

Real-time quantitative PCR (qPCR) is the most widely used method for DNA quantification.[112] Qubit double stranded DNA (dsDNA) quantification assays (Thermo Fisher Scientific) use fluorescent dyes that selectively intercalate with dsDNA molecules. Qubit fluorometry is a very fast and cost-effective tool to measure DNA and proved to yield reliable results in low concentrated samples and cirDNA-samples (<0.1ng/µl).[112,113]

1.5.3 Other substrates of liquid biopsies

Circulating tumor cells (CTCs) are found in the blood stream of cancer patients in extremely low numbers, yet different isolation methods can allow to characterize single

CTCs. As opposed to ctDNA not only genomic investigations can be conducted but also the transcriptome, proteome and secretome can be studied.[114]

Tumor-educated platelets (TEPs) influence the systemic and local response to malignant tumors and are capable of priming tumor cells for dissemination. By profiling the transcriptome of platelets, tumor patients can accurately be identified, and the tumor entity determined. Even deductions about the mutation status of common hotspots, like KRAS and EGFR can be made.[61,115]

The methylome of cirDNA is a sensitive approach to identify low levels of ctDNA and enables tumor classification.[116,117] Epigenetic alterations occur in all cancers and epigenetic silencing of MLH1 is the most common cause for MSI in sporadic colorectal cancer.[39]

Investigating different entities of circulating nucleic acids (cirNA) may yield complementary information, as for example exosomes might be secreted by a different subset of cells or tumor cells.[80,118] Analysis of cirRNA expression profiles, either messenger RNAs or micro RNAs, can give provide information about the tumor dynamics and its origin.[119]

1.5.4 Applications of liquid biopsy in cancer

In cancer patients cirDNA-levels were confirmed to have an independent prognostic effect.[120] Strikingly, higher levels or fractions of ctDNA have consistently been associated with worsening overall survival[92,121,122] and patients without detectable tumor DNA had a better prognosis.[123,124]
Serial liquid biopsies during treatment are convenient as they can be added to routine phlebotomy. ctDNA dynamics correlate with clinical response[99,122,125,126] and resemble the overall tumor burden.[100] In fact, ctDNA monitoring can recognize response earlier than medical imaging or serum biomarkers.[97,127] By screening for actionable hotspot mutations, emerging therapy resistance can be discovered,[92,97,124] hereby offering physicians the needed information to discontinue a treatment and protect the patient from unnecessary adverse effects. This highlights its predictive value in monitoring patients receiving palliative chemotherapy regimens or targeted therapies.

Monitoring of minimal residual disease (MRD) and recognition of recurrence in patients with early stages of disease requires extremely sensitive detection methods. In CRC

possible surveillance programs, including imaging studies and endoscopic procedures, have limited usability[68] and a meta-analysis showed no improved overall survival through intensified follow-up.[128]

In a large cohort of patients with stage II CRC the postoperative presence of ctDNA could reliably indicate MRD and those patients had recurrence in the next three years, whereas the 3-year recurrence-free survival in ctDNA-negative patients reached 90%.[129] Other reports strengthen the relevance of ctDNA in early-stage cancer and recurrence detection.[99,130,131] cirDNA analysis has the potential to be a screening tool for groups or symptomatic patients. ctDNA in individuals with no known tumor has been detected up to two years before these individuals were diagnosed with cancer.[132] In patients with cancer of unknown primary (CUP-syndrome) liquid biopsy can help classify the tumor origin and consequently allow adequate treatment to be initialized promptly.[133]

1.6 Cancer genomics and its implications for liquid biopsies

The knowledge from thousands of individual cancer genomes and functional genomic studies have characterized the biology of many genes and the impact of structural aberrations on their functions. Nonetheless, there is the need for genomics data of tumors under therapy, to understand the impact of genomic alterations in tumor progression and drug response.

1.6.1 Next-generation sequencing (NGS)

Different NGS technologies have been developed, among them the Illumina platforms. Here, DNA templates are prepared by fragmentation into pieces of 200 bases. Then, specific adapters bind to both ends, enabling hybridization to a flow-cell. Bridge amplification is performed, generating clusters of up to a million copies of each template on the flow-cell surface. Sequencing by synthesis is the last step. This step is very similar to Sanger sequencing with the difference that the Illumina based methods not only sequence one template but millions simultaneously. Fluorescence labelled nucleotides are added each cycle and each time an image of the clusters is obtained. These pictures can be ordered and analyzed, determining the nucleotide sequence of each cluster.[134] Using bioinformatic tools, the individual reads are mapped against a reference genome and variants can be called.

Limiting factors of Illumina platforms are the read length (maximum of 300bp until now) and the use of multiple PCR cycles, as thereby errors are introduced.[135]

1.6.2 Approaches to ctDNA analysis

Depending on the initial question different sequencing techniques can focus on specific point mutations or analyze larger parts of the genome, even the whole genome.
Assays for single-locus investigations can detect even individual mutant copies per ml.[77] Those assays are often based on digital droplet PCR (ddPCR). ddPCR is a refinement of basic PCR techniques. Through water-in-oil emulsion samples are divided into smaller reactions and then oil-droplets are generated. Consequently, thousands of separate PCR reactions are performed. As ddPCR amplifies single templates in each reaction, the DNA input can be quantified in an absolute manner.[136] Other single-locus techniques include BEAMing, Intplex and ARMS-PCR.[77] First commercial tests are now implemented in clinical diagnostics.[137]
To interrogate greater numbers of hotspots and larger regions NGS-based approaches are needed. Variable sequencing scales can be used ranging from whole-genome (WGS) and whole-exome (WES) to targeted sequencing (amplicon-based or hybrid-capture). Amplicon panels target only specific genomic loci and use different enrichments strategies.[138] Specific primers flank the region of interest, followed by amplification of the respective amplicon. Further sample preparation is performed by attaching sequencing adapters either by ligation or by nested PCR to the amplicons. Amplicon sequencing can monitor clonal evolution and disease burden under treatment without prior knowledge of mutational statuses. The detection limit for different targeted technologies ranges between 0.01% and 1% MAF.[77]
WES and WGS are still very expensive and the current detection limit of 5-10% MAF reduce their application field to copy number alterations or aneuploidies and cases of high-tumor burden.[77]

1.7 Aims of this thesis

cirDNA from blood samples is an easily accessible source of tumor material representing all tumor sites.[92,139,140]
Several areas of oncology would profit from novel diagnostic markers. Current biomarkers have a low sensitivity and specificity, frequent radiological scans expose patients to ionizing radiation and tissue biopsies are often not feasible.

Monitoring ctDNA by sequencing of frequent or even actionable alterations may indicate tumor progress or acquired resistance to therapy even before radiologic progress.[92,96,97] Therefore cirDNA analysis has can potentially guide personalized medicine and complement or even replace prevailing diagnostic investigations.
To be implemented into daily routine, the analysis of cirDNA needs to be fast and cost-efficient.

Hence, the aims of this thesis can be divided into two parts:

a. Clinical relevance of cirDNA-levels in patients with gastrointestinal tumors:
   cirDNA-levels are compared to clinical parameters as well as blood values, including serum tumor biomarkers.

b. Establishment and assessment of a customized amplicon panel for the analysis of cirDNA in patients with mCRC:
   An experimental and bioinformatic pipeline for amplicon sequencing is established for the analysis of cirDNA-samples. A panel targeting frequent and potentially actionable hotspots in CRC is designed. The panel is then applied to a clinical cohort of mCRC with multiple blood samples over time to show the relevance of cirDNA analysis compared to imaging and serum biomarkers.

2 MATERIALS AND METHODS

2.1 Materials

Reagents and Chemicals

Chemicals, Reagents, Markers	Supplier
6*DNA loading dye	Thermo Fisher Scientific
Acetic Acid	Sigma-Aldrich
Agarose	Sigma-Aldrich
AmpureXP beads	Beckman Coulter
dNTPs	Thermo Fisher Scientific
EDTA	Sigma-Aldrich
Ethanol absolute	VWR Chemicals
Glycerol	Sigma-Aldrich
Gene Ruler 100bp	Thermo Fisher Scientific
KHCO3	Merck
NH4Cl	Merck
PBS	Sigma-Aldrich
Q5 Hot Start High-Fidelity buffer	New England BioLabs
Q5 Hot Start High-Fidelity DNA Polymerase	New England BioLabs
sodium citrate	Fluka
sodium phosphate dibasic	Roth
sodium phosphate monobasic, monohydrate	Merck
SYBR Safe DNA gel stain	Thermo Fisher Scientific
Tris base	Roth
Xylene	MP Biomedicals

Kits

Kits and assays	Supplier
Agilent DNA 1000 Kit	Agilent Technologies
Agilent High Sensitivity DNA kit	Agilent Technologies
DNeasy Blood and Tissue kit	Qiagen
MiSeq Reagent kit V2	Illumina
QIAamp DNA FFPE kit	Qiagen
QIAamp DNA Blood Mini kit	Qiagen
Qubit dsDNA HS Assay Kit	Thermo Fisher Scientific

Buffers

Buffer	Ingredients
10x Red Blood Cell Lysis Buffer	8.3g NH4Cl

	1g KHCO3 200µl EDTA 0.5 M 100ml H20
Storage buffer	99% glycerol 50nM sodium citrate 20mM sodium phosphate monobasic, monohydrate 20mM sodium phosphate dibasic
TAE 50x stock solution	242.0g/l Tris base 57.1ml/l Acetic acid 100ml/l EDTA 0.5M

Cell culture media and supplements

Cell culture media and supplements	Supplier
0.25% trypsin	Sigma-Aldrich
Fetal bovine serum (FBS)	PAA (Thermo Fisher Scientific)
McCoy's 5A medium	Gibco (Thermo Fisher Scientific)
PBS	Sigma-Aldrich

Cell lines

Cell line	Source	Medium	Provided by
HT29	ATCC	McCoys 5A with 10% FBS	ATCC
HCT116	ATCC	McCoys 5A with 10% FBS	ATCC

Primers first PCR

Universal adapter sequence (UAS)	Added to
TCCCTACACGACGCTCTTCCGATCT	5' end of each FW primer
AGTTCAGACGTGTGCTCTTCCGATC	5' end of each REV primer

Gene	targeted amplicon	FW or REV	Sequence (without UAS)
APC	APC_1	FW	CAGAGTAGAAGTGGTCAGCC
		REV	CCTGTGTCGTCTGATTACATCC
APC	APC_2	FW	TTGGAATATGAAGCAAGGCA
		REV	TCGCTGTTTTATCACTTAGAAAC
APC	APC_3	FW	CAGGTTATTGCGAGTGTTTTGAG
		REV	ATTCCATCAATGCTTTCACACTTCC
APC	APC_4	FW	GAGAGAACGCGGAATTGGTC
		REV	ATGACTTTGGCAATCTGGGC

APC	APC_5	FW	TGGACAGCAGGAATGTGTTTC
		REV	TCTTCTTGACACAAAGACTGGC
APC	APC_6	FW	AAGCCCCAGTGATCTTCCAG
		REV	AGCTTGCTTAGGTCCACTCT
APC	APC_7	FW	GGTGCTCAGACACCCAAAAG
		REV	GCCACTTACCATTCCACTGC
APC	APC_8	FW	ATCATCTTTGTCATCAGCTGA
		REV	TGGAACTTCGCTCACAGGAT
BRAF	BRAF_1	FW	GCTCTGATAGGAAAATGAGATCTACTG
		REV	CCATCCACAAAATGGATCCAGACA
DPYD	DPYD_1	FW	GTGTAGAAATGGCCGGATTGAAG
		REV	GAGAAAGTTTTGGTGAGGGCAAA
DPYD	DPYD_2	FW	TGTCTCAGTGTTTGCCCTATTGT
		REV	ATCACCTTAACACACCGGATTCA
ERBB3	ERBB3_1	FW	GGATTCGAGAAGTGACAGGC
		REV	GCAAACTTCCCATCGTAGACC
FBXW7	FBXW7_1	FW	TCATCATTAGTGGATCTACAG
		REV	CAATAATAGAGGAAGAAGTCC
FBXW7	FBXW7_2	FW	TCTGCAGAGTTGTTAGCGGT
		REV	AACAACCCTCCTGCCATCAT
GNAS	GNAS_1	FW	CTGTTTCGGTTGGCTTTGGT
		REV	TGGAAGTTGACTTTGTCCACC
KRAS	KRAS_1	FW	ATTATAAGGCCTGCTGAAAATGACTGA
		REV	GGTCCTGCACCAGTAATATGC
KRAS	KRAS_2	FW	ATCCAGACTGTGTTTCTCCCTTC
		REV	GGCAAATACACAAAGAAAGCCCT
MLH1	MLH1_1	FW	GAGAGACAGTAGCTGATGTTA
		REV	AATGTGATGGAATGATAAAC
MLH1	MLH1_2	FW	GAAAAATCAATCTTCTGTTC
		REV	CTAGAACACATTACTTTGATGA
MLH1	MLH1_3	FW	GACCTCGTCTTCTACTTCTGGAA
		REV	CTTATCCTCTGTGACAATGGCCT
MLH3	MLH3_1	FW	CTACTGAAGTGGGATGCCAGC
		REV	TGGAACATAATTTAACTCGCCCA
MSH6	MSH6_1	FW	TTAACAGATGTTTTACTGTGC
		REV	TAGGAATAAAATCATCTCCA
NRAS	NRAS_1	FW	ACAAACTGGTGGTGGTTGGA
		REV	CACTGGGCCTCACCTCTATG
NRAS	NRAS_2	FW	CACACCCCCAGGATTCTTAC
		REV	TGGCAAATACACAGAGGAAGC
PIK3CA	PIK3CA_1	FW	ACTCAAGAAGCAGAAAGGGAAG
		REV	CGGTTGCCTACTGGTTCAAT

PIK3CA	PIK3CA_2	FW	CAGAGTAACAGACTAGCTAGAGACA
		REV	AGCACTTACCTGTGACTCCA
PIK3CA	PIK3CA_3	FW	TAGGCAAGTCGAGGCAATGG
		REV	GGCTTCTAAACAACTCTGCCC
PIK3CA	PIK3CA_4	FW	TGCATACATTCGAAAGACCCT
		REV	GCATGCTGTTTAATTGTGTGGAAG
PMS2	PMS2_1	FW	GCCATTCAAACCAGGAAGATACC
		REV	GAGGCTGACATGTCCTGAGTATT
PTEN	PTEN_1	FW	TGTGAAGATCTTGACCAATGGCT
		REV	TAGAAATCTAGGGCCTCTTGTGC
SMAD2	SMAD2_1	FW	CCTCTACAGTGGTTGGACAAAGT
		REV	GACCACACACAATGCTATGACAG
SMAD3	SMAD3_1	FW	CATCTCCTACTACGAGCTGAACC
		REV	CTGCATTCCTGTTGACATTGGAG
SMAD4	SMAD4_1	FW	TAGGAGAGACATTTAAGGTT
		REV	ACTATACAATCAATACCTTGC
SMAD4	SMAD4_2	FW	AAATTCACTTACACCGGGCC
		REV	TAAGGTTAAGGGCCCCAACG
TGFBR2	TGFBR2_1	FW	ATAACACTAGAGACAGTTTGC
		REV	TTCTGAGAAGATGATGTTG
TGFBR2	TGFBR2_2	FW	TGTGTGAGACGTTGACTGAGTG
		REV	ATTTGGTAGTGTTTAGGGAGCCG
TP53	TP53_1	FW	GGACCTGATTTCCTTACTGCC
		REV	CTCCCCTTTCTTGCGGAGA
TP53	TP53_2	FW	TTGGCTCTGACTGTACCACC
		REV	CAAGTGGCTCCTGACCTGG
TP53	TP53_3	FW	TGGCCATCTACAAGCAGTCA
		REV	TCAGTGAGGAATCAGAGGCC
TP53	TP53_4	FW	GGCCTCTGATTCCTCACTGA
		REV	CCAGTTGCAAACCAGACCTC
TYMP	TYMP_1	FW	ATACCAGGGGCCATGCTGAT
		REV	CCCTGTGGAATGCTTGTCCA
TYMS	TYMS_1	FW	CCAAGGGAGTGAAAATCTGGGAT
		REV	GGCATTGTTCTATCTCCTCACCT
UPP1	UPP1_1	FW	GTGTGTCACCCTCCTGAACC
		REV	AATGGCAAGTCATCACAGCAGG

Primers second PCR

Primer	**Sequence** (with respective TruSeq DNA HT indexes)
Forward primer	AATGATACGGCGACCACCGAGATCTACAC-(Index)-ACACTCTTTCCCTACACGACGCTCTTCCGATCT
Reverse primer	CAAGCAGAAGACGGCATACGAGAT-(Index)-GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC

All primers were produced by Eurofins Genomics. (Louisville, USA)

Consumables

Article	**Supplier**
5ml self-standing conical tubes	Axygen Scientific
10µl, 20µl, 200µl and 1000µl pipet tips	Starlab/Gilson
15ml and 50ml conical tube	Thermo Fisher Scientific
96-well full-skirted plates	Thermo Fisher Scientific
96-well hard-shell PCR plates skirted black	Bio-Rad
Aluminium foil lids	Beckman Coulter
Biomek AP96 P250 pipet tips	Beckman Coulter
Biomek AP96 P20 pipet tips	Beckman Coulter
Cryovials 2ml Cryo.s	Greiner Bio-One
Eppendorf tube 1.5ml	Eppendorf
Eppendorf tube 2ml	Eppendorf
Falcon serological pipettes	Thermo Fisher Scientific
Finntip 250 Universal pipet tips	Thermo Fisher Scientific

Instruments

Instrument	**Manufacturer**
Bioanalyzer 2100	Agilent Technologies
Biomek FX	Beckman Coulter
Biometra thermal cycler T advanced	Analytik Jena
Bioshake IQ	BioShake
Centrifuge 5424	Eppendorf
Centrifuge 5804	Eppendorf
Centrifuge 5804 R	Eppendorf
Centrifuge Rotanta/RP	Hettich
Centrifuge Sigma 3K12	Sigma-Aldrich
DNAEnging thermal cycler	BioRad
E-Box VX2	Vilber
Mini Sub Cell GT	Bio-Rad
MiSeq	Illumina

NanoDrop ND-1000	Marshall Scientific
PowerPacUniversal	Bio-Rad
Qubit Fluorometer 2.0	Thermo Fisher Scientific
Vortex Genie 2	Scientific Industries
Wide Mini Sub Cell GT	Bio-Rad

Software

Software	**Supplier**
BLAST+	NCBI
Excel, PowerPoint, Word for Office 365 ProPlus v. 1903	Microsoft
FASTQC v. 0.11.6	Babraham
GraphPad Prism 8.1.0	Graph Pad
Illustrator CC v. 17.1	Adobe
MiSeq Reporter v2.6	Illumina
MultiPLX 2.1	ELIXIR; http://bioinfo.ut.ee/multiplx/
primer3	ELIXIR; http://primer3.ut.ee/
RStudio v. 1.1.419	RStudio
SAM tools v. 1,4	Li et al.

2.2 Clinical patients

2.2.1 Patient recruitment

All patients with gastrointestinal tumors receiving treatment at the *Tagestherapiezentrum* or those who were implemented into an aftercare program in the outpatient clinic of the University Hospital Mannheim were eligible for study inclusion. In addition, patients receiving stationary treatment were screened and asked for participation, if they met the inclusion criterion of having an underlying gastrointestinal tumor. Written informed consent was obtained from all patients.

2.2.2 Ethics board approval

Ethics board approval was obtained from the local ethics committee (Ethikkomission II, Medical Faculty Mannheim, Heidelberg University, identifier 2013-640N-MA).

2.2.3 Blood collection

Upon regular blood withdrawals extra tubes were taken for the study purpose. If possible, samples were collected before start of a new treatment and every time a patient needed inpatient care. Otherwise standardized sample collection every 3 to 6 months during radiological re-staging was performed.
Simultaneously, the following blood values were determined:
Albumin, bilirubin, creatinine, ALAT, ASAT, GGT, ALP, LDH, WBC count, hemoglobin level, platelet count, INR, CRP and serum tumor markers CA19-9, CEA and AFP.

2.2.4 Documentation of clinical characteristics

Patient data were stored in a pseudonymized fashion in the prospective MALIBU clinical database. Relevant patient characteristics and clinical information were recorded.

Patient characteristics:
Age, sex, history of tumor disease, underlying inflammatory bowel disease or liver disease.

Tumor characteristics:
The following information were collected for every patient at the time of blood sampling: Tumor entity, localization, histology, time of diagnosis, TNM classification and UICC stage at initial diagnosis and at the time of the first sample collection (or BCLC stage for HCC and Bismuth classification for patients with hilar cholangiocarcinoma), metastatic sites and molecular markers (microsatellite status, KRAS, NRAS and BRAF mutational status, ERBB2 expression). The latter molecular markers were routinely determined by the Institute of Pathology, University Hospital Mannheim from tumor biopsy tissue. Sanger sequencing was routinely performed for DNA mutation detection by the Institute of Pathology.

Therapy history:
Current therapy, treatment setting (surveillance, neoadjuvant, adjuvant or palliative) and line of treatment, previous systemic treatments as well as radiation, interventional and surgical therapies.

Along with new samples the aforenamed blood values and the clinical re-staging results were documented. Changes in therapy regimes were noted.

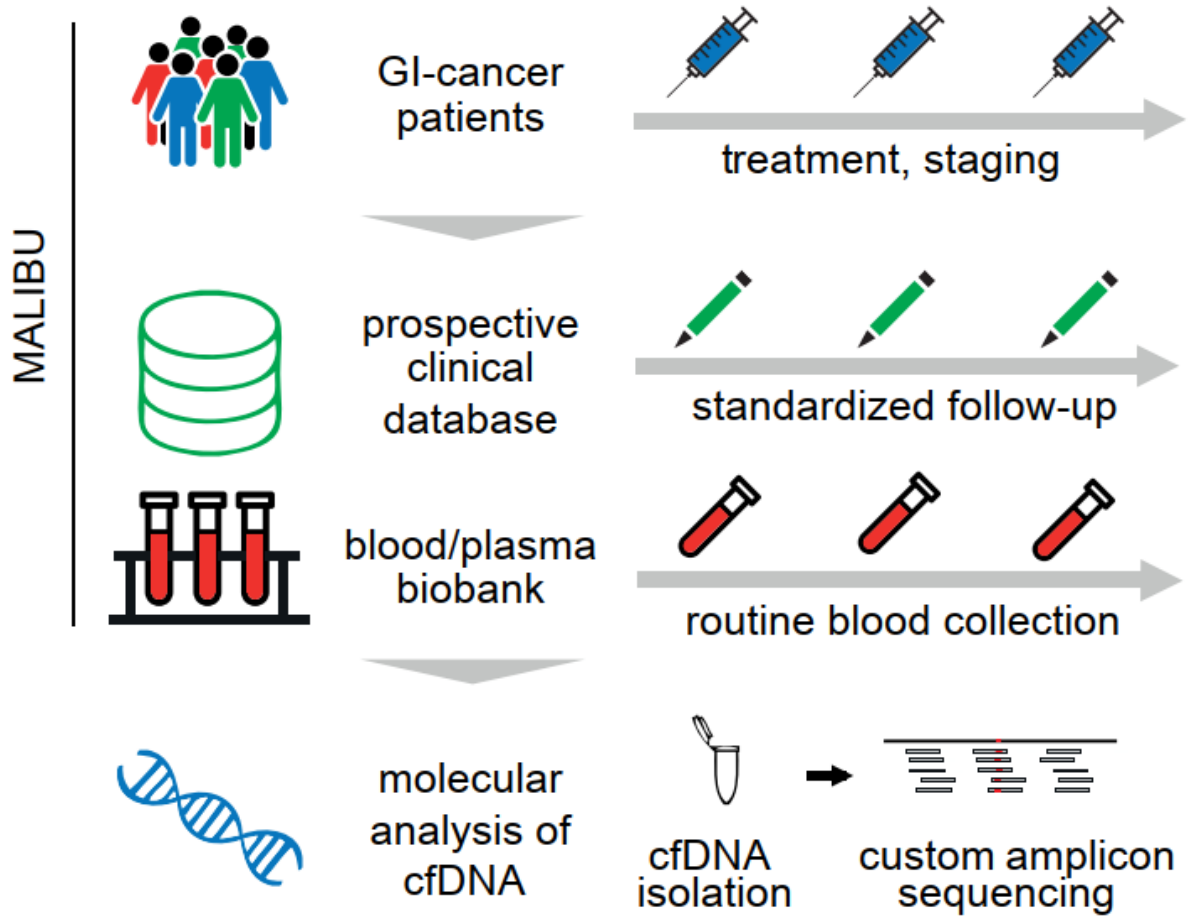

Figure 2-1. Standardized collection of blood plasma and clinical data. The Mannheim Liquid Biopsy Unit (MALIBU) prospectively collects blood samples as well as clinical data from patients with gastrointestinal tumors. Blood is collected at regular time intervals over the course of treatment and follow-up. The plasma is used for isolation of cirDNA and subsequently for custom amplicon sequencing.

2.2.5 Patients and cohort selection

Between January 2015 and October 2017 280 patients treated for a gastrointestinal tumor at the University Hospital Mannheim were included in the study. 972 blood samples were collected during this time span. A total of 389 blood samples were further processed for cirDNA analysis.

To investigate the role of cirDNA in advanced gastrointestinal tumors (GI-tumors), a group of 34 patients with metastasized or advanced CRC, from now on referred to as mCRC-cohort, were selected. All patients received a palliative chemotherapy with a

5-FU backbone. Initially all patients showed a therapy response as confirmed by imaging studies and the RECIST 1.1 criteria.[141] At least two samples were collected during therapy. Characteristics of the study cohorts are depicted in Table 2-1 and Table 2-2.

From 21 CRC patients (19 belong to the mCRC-cohort) tumor tissue in form of formalin fixed and paraffin embedded (FFPE) samples - either primary tumor or metastasis - was obtained from the archives of the Institute of Pathology, University Hospital Mannheim (Table 2-3).

Table 2-1. MALIBU study: patient, treatment and tumor baseline characteristics. *

Parameter	***All patients n=280***
Gender	
Male	193 (69%)
Female	87 (31%)
Age (years)	66 (32-92)
Tumor entity	
CCC	24 (9%)
CRC	104 (37%)
esophagus cancer	36 (13%)
gastric cancer	27 (10%)
GIST	1 (<0.5%)
HCC	31 (11%)
NET	12 (4%)
pancreatic cancer	44 (16%)
small intestine cancer	1 (<0.5%)
multiple entities	8 (3%)
Stage**	
I	21 (9%)
II	21 (9%)
III	55 (23%)
IV	145 (60%)
Therapy setting	
surveillance	37 (13%)
curative	52 (19%)
palliative	191 (68%)
Metastases	
hepatic	98 (63%)
pulmonary	50 (32%)
peritoneal carcinomatosis	46 (30%)
osseous	17 (11%)
pleural carcinomatosis	8 (5%)
other	5 (8%)

*n is shown for categorical variables with percentage in parentheses.
For continuous variables, median is shown with range in parentheses.

** without HCC and Klatskin-tumors.

Table 2-2. CRC-cohort: patient, treatment and tumor baseline characteristics. *

Parameter	***All CRC n*=104**	***mCRC-cohort n*=34**
Gender		
Male	74 (71%)	25 (74%)
Female	30 (29%)	9 (26%)
Age (years)	65 (34-88)	65 (36-88)
Stage		
I	9 (9%)	
II	10 (10%)	
III	25 (24%)	3 (9%)**
IV	60 (58%)	31 (91%)
Tumor location		
colon cancer	45 (43%)	15 (44%)
rectal cancer	58 (56%)	19 (56%)
both entities	1 (1%)	
Therapy setting		
surveillance	24 (23%)	
adjuvant/neoadjuvant	21 (20%)	
palliative	59 (57%)	34 (100%)
Metastases (only stage IV)		
hepatic	47 (78%)	24 (71%)
pulmonary	28 (47%)	18 (53%)
peritoneal carcinomatosis	14 (23%)	4 (12%)
pleural carcinomatosis	5 (8%)	2 (6%)
osseous	4 (7%)	1 (3%)
other	5 (8%)	0 (0%)
Therapy regimen		
5-FU	17 (16%)	6 (18%)
FOLFOX/CAPOX	18 (17%)	4 (12%)
FOLFIRI	37 (36%)	22 (65%)
FOLFOXIRI	3 (3%)	2 (6%)
irinotecan	2 (2%)	0 (0%)
no chemotherapy	25 (24%)	0 (0%)
with anti-VEGFR-antibody	25 (24%)	16 (47%)
with anti-EGFR-antibody	27 (26%)	13 (38%)

*n is shown for categorical variables with percentage in parentheses.
For continuous variables, median is shown with range in parentheses.

** these three patients all had an inoperable relapse of their primaries.

Table 2-3. Patients with tumor tissue undergoing amplicon sequencing.

Patient No.	Sex	Age	Tumor	tissue origin	ng DNA	mCRC-cohort
0019	F	76	rectum cancer	primary	413	yes
0024	F	55	colon cancer	metastasis	100	yes
0025	M	62	rectum cancer	metastasis	620	yes
0044	M	86	colon cancer	primary	47.3	yes
0055	M	51	rectum cancer	primary	17.3	yes
0065	F	49	rectum cancer	metastasis	281	no
0080	M	78	colon cancer	primary	565	yes
0083	M	72	rectum cancer	primary	3390	yes
0109	M	66	rectum cancer	primary	38.1	yes
0119	F	68	rectum cancer	primary	1160	yes
0133	F	59	rectum cancer	primary	545	yes
0143	M	59	rectum cancer	primary	128	yes
0144	M	80	colon cancer	primary	1600	yes
0149	M	59	colon cancer	primary	745	yes
0152	F	36	rectum cancer	primary	780	yes
0167	M	59	colon cancer	primary	3470	yes
0168	M	73	rectum cancer	primary	1358	yes
0324	F	67	rectum cancer	primary	1555	yes
0340	M	57	colon cancer	metastasis	19.7	no
0427	M	44	colon cancer	primary	384	yes
0430	F	57	colon cancer	metastasis	66	yes

2.3 Sample preparation and DNA isolation

2.3.1 Plasma and buffy coat preparation

After collection of whole blood in two blood collecting tubes containing EDTA as anticoagulant (9ml EDTA-tube, Sarstedt, Nürmbrecht, Germany), samples were stored at 4°C and further processed within 16 h. Sample preparation was adapted to earlier descriptions and in line with the recommendations of a dedicated review.[99,108]

Samples were first centrifuged at 4°C and 1900*g for 10 min with the brake off, thereby separating plasma from cells. Five mm above the buffy coat, the supernatant, representing the plasma, was carefully transferred into a fresh 15ml falcon tube. The plasma of two blood collection tubes was combined.

Pooled plasma was distributed into 2ml Eppendorf tubes (Eppendorf, Hamburg, Germany) and again centrifuged, now at 4°C and 16000g for 10 min in a fixed-angle

rotor. Without dislodging the formed cell pellet, plasma was transferred into 2ml cryovials (Greiner Bio-One, Frickenhausen, Germany) and stored at -80°C until needed. That way contamination by cellular material was minimized.
Additionally, the buffy coat was conserved. After the plasma had been separated, the concentrated leukocyte band, i.e. buffy coat, was absorbed by pipetting and transferred into a 15ml centrifugation tube. Five times the volume of red blood cell (RBC) lysis buffer was added, the tube inverted a few times and left 10 min at room temperature until the liquid was clear red. PBS was added up to 15ml and centrifuged for 10 min at 250g. Supernatant was decanted, the cell pellet resuspended in 5ml of PBS and spun down as above. Again, the supernatant was decanted, and the tubes inverted to drain them briefly. In the end the cell pellet was resuspended in 1ml of cold storage buffer and stored at -80°C.

2.3.2 cirDNA-isolation and quantification

For cirDNA isolation, 1-4ml of frozen plasma was used. Extraction of DNA was performed using the QIAamp DNA Blood Mini Kit (Qiagen, Hilden, Germany). The manufacturer's vacuum protocol using the QIAvac 24 Plus was applied. In the end DNA was dissolved in 50-70µl nuclease-free water and stored at -20°C for further use. The amount of DNA was quantified with Qubit 2.0 fluorometer (Thermo Fisher Scientific, Waltham, USA), using the high sensitivity assay according to the manufacturer's protocol and 3µl of each sample. If no DNA was detected, measuring was repeated with 10µl.

2.3.3 FFPE-samples and cell lines

FFPE-samples

Tissue, either from the primary tumor or a hepatic metastasis, was fixed in formalin and then embedded in paraffin as part of routine pathological management. Qualified blocks were chosen and three 10µm slices were used for DNA extraction.
A corresponding hematoxylin and eosin stained slide was used to estimate the tumor cell content. This assessment was performed by Prof. T. Gaiser, an expert pathologist. By macrodissection the tumor DNA content was enriched.
DNA was isolated using QIAmp DNA FFPE kit (Qiagen) according to the manufacturer's instructions. DNA was eluted in 50µl and concentrations were

measured using the Qubit 2.0 fluorometer. Isolation yielded between 17.3ng (small metastasis sample) and 3470ng (primary tumor sample) of DNA.

Cell lines

All cell lines were obtained from ATCC. HCT116 and HT29 cells were maintained in McCoy's 5A medium (GIBCO), both with 10% FBS (PAA). No antibiotics were used. Cells were cultured at 37°C in humidified atmosphere with 5% CO2. Cells were passaged every 2-3 days using 0.25% trypsin (Sigma-Aldrich, St Louis, USA).

Cells from the tumor cell lines were pelleted and DNA isolation was performed with the DNeasy Blood & Tissue kit (Qiagen) according to the manufacturer's recommendations. Extracted DNA was diluted and stored at -20°C for further use. DNA concentrations were measured using the NanoDrop ND-1000 (Marshall Scientific, Hampton, USA).

2.4 Amplicon sequencing of cirDNA

2.4.1 Design of amplicon panel

Positions of frequently mutated hotspots in CRC associated genes were obtained either from the COSMIC database or cbioportal.[33,142] Amplicons covering mutations related to emerging resistance against anti-EGFR antibodies were designed.[52] In addition, literature search was performed for genes and mutations related to 5-FU resistance. Functional regions of genes that regulate the metabolism of 5-FU (DPYD, TYMS, TYMP, UPP1)[70,71,73] were identified. In addition, genes relevant for mismatch repair were included.[33,39] Alterations in these genes occur at low frequencies and are not commonly covered by commercial amplicon panels. Together they cover a large mutational spectrum in CRC. An overview of all genes is visualized in Figure 2-2.

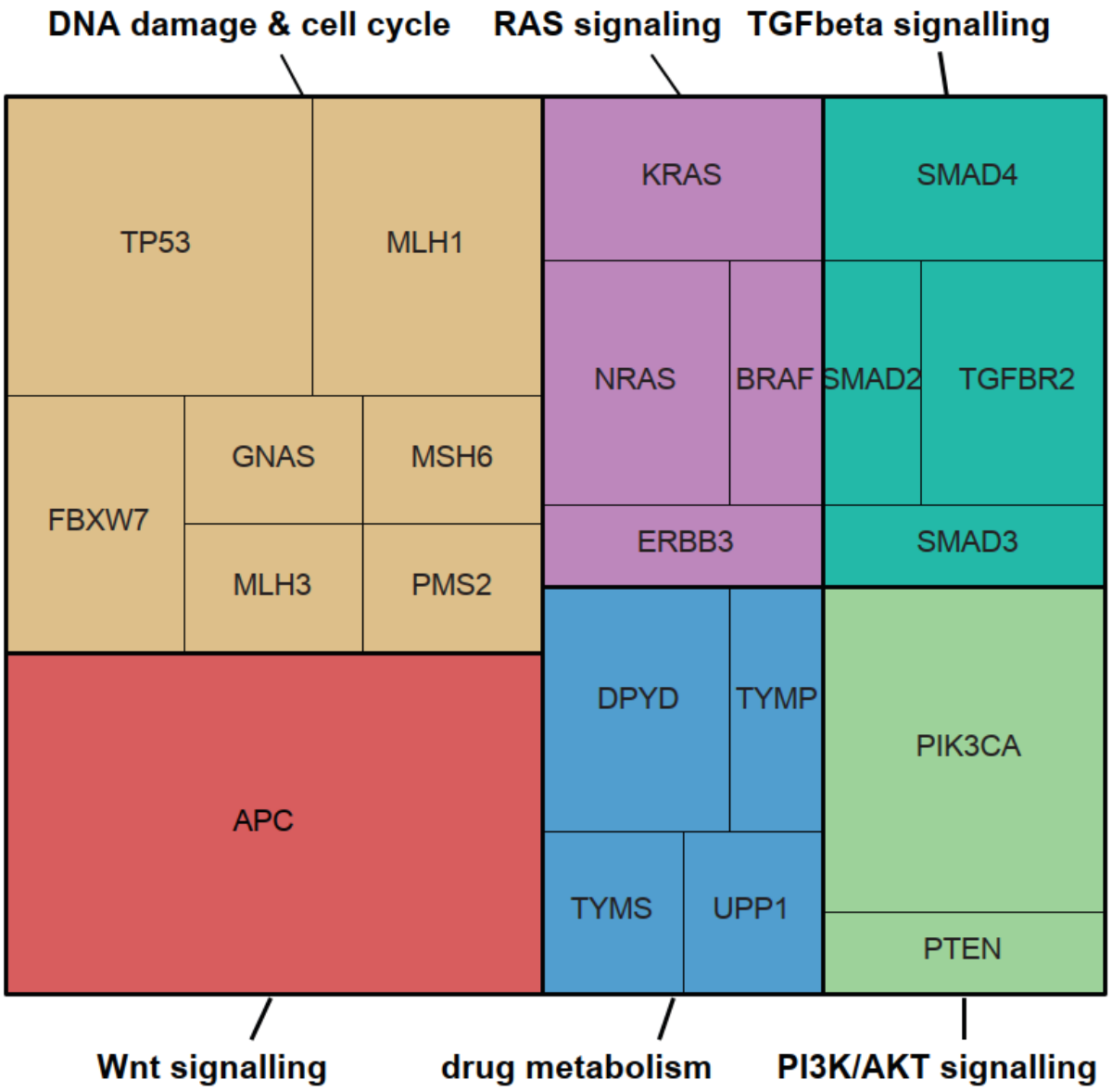

Figure 2-2. Tree map of genes included in the custom amplicon panel. A total number of 43 amplicons was designed to detect hotspot mutations in 18 genes frequently altered in colorectal cancer, including genes related to DNA damage and cell cycle, RAS signaling, TGF beta signaling, Wnt signaling, PI3K/AKT signaling, as well as four genes related to drug metabolism.

2.4.2 Design of primers and multiplex reaction

Nucleotide sequences surrounding the genetic alteration of interest were obtained from the National Center for Biotechnology Information (NCBI) database with the genome assembly GRCH38 as a reference. Specific primers for regions that cover the selected mutations were designed with primer3 with the following settings:[143]

Primer bind ~50bp up-and downstream of region of interest, primer length of 19 to 27bp, GC ratio of 30 to 63% and melting temperature of 55 to 61.5°C.

The universal adapter sequence (UAS) TCCCTACACGACGCTCTTCCGATCT was added to the 5' end of each forward primer. To each reverse primer the sequence AGTTCAGACGTGTGCTCTTCCGATC was added to the 5' end. The amplicon length varied between 100 to 175bp. Sequences of all primers can be found in Chapter 2.1 and further amplicon and primer characteristics in Table 2-4.

Agarose gel electrophoresis

TAE 50*stock solution was diluted 1:49 with water to produce TAE-buffer. 2g agarose per 100ml TAE-buffer were used to produce 2% agarose gels. 2% concentration is better suited to discriminate DNA fragments smaller than 500bp. After heating all agarose was dissolved in the TAE-buffer and was poured into the gel slots. 10-15µl SYBR Safe DNA gel stain (Thermo Fisher Scientific) was added, depending on the gel size. After approximately 30 min, the gel had polymerized.

DNA ladder, controls and respective samples could now be filled into the pockets. 2-5µl loading dye was added to each sample. Applying 125mV, the DNA fragments were separated by size in 15-20 min. DNA fragments were detected afterwards using the E-Box VX2 (Vilber, Collégien, France).

Primer pairs amplifying predominantly primer dimers seen as fragments ~100bp were discarded and redesigned. In the end, primers were validated by sequencing.

To develop a multiplex PCR assay, the software tool MultiPLX 2.1 was used to identify appropriate primer combinations.[144] To this end, the concentration of monovalent salts was set to 50mM and the concentration ofmg2+ to 1.5mM. All five possible primer interaction scores that might affect primer compatibility were calculated. To group the 43 primer pairs for multiplex PCR the stringency value "normal" was applied. The best combination resulted in six different pools with 5-9 primer pairs.

Table 2-4. Amplicon panel. 43 primer pairs were designed for hotspots of 18 genes and multiplexed in six reactions. The most frequent mutation of the hotspot is listed. Primer length and GC-content are named for each primer.

Amplicon	Hotspot	Primer	Primer length	GC%	Multiplex reaction
APC_1	1245-1273	FW	20	55	5
		REV	22	50	
APC_2	R213*	FW	20	40	6
		REV	23	35	
APC_3	R564*	FW	23	43	6
		REV	25	40	
APC_4	R876*	FW	20	55	6
		REV	20	50	
APC_5	R1114*	FW	21	48	5
		REV	22	45	
APC_6	R1450*	FW	20	55	5
		REV	20	50	
APC_7	Q1378*, Q1367	FW	20	55	5
		REV	20	55	
APC_8	Q1294*	FW	21	38	5
		REV	20	50	
BRAF_1	V600E	FW	27	41	1
		REV	24	46	
DPYD_1	R561Q	FW	23	48	1
		REV	23	43	
DPYD_2	P1010H	FW	23	43	2
		REV	23	43	
ERBB3_1	P104	FW	20	55	1
		REV	21	52	
FBXW7_1	R465	FW	21	38	2
		REV	21	38	
FBXW7_2	R505, R479	FW	20	50	3
		REV	20	50	
GNAS_1	R201	FW	20	50	3
		REV	21	48	
KRAS_1	G12	FW	27	37	4
		REV	21	52	
KRAS_2	A59T, Q61H	FW	23	48	2
		REV	23	43	
MLH1_1	R226, I219V	FW	21	43	4
		REV	20	30	
MLH1_2	N158	FW	20	30	4
		REV	22	32	
MLH1_3	V384D	FW	23	48	3
		REV	23	48	
MLH3_1	E586, S587	FW	21	57	3
		REV	23	39	
MSH6_1	F1088	FW	21	33	3
		REV	20	30	
NRAS_1	G12, G13	FW	20	50	4
		REV	20	60	

Amplicon	Hotspot	Primer	Primer length	GC%	Multiplex reaction
NRAS_2	Q61	FW	20	55	1
		REV	21	48	
PIK3CA_1	R88	FW	22	45	4
		REV	20	50	
PIK3CA_2	E545, E542, Q546	FW	25	44	3
		REV	20	50	
PIK3CA_3	E726	FW	20	55	4
		REV	21	52	
PIK3CA_4	H1047	FW	21	43	3
		REV	24	42	
PMS2_1	R563	FW	23	48	4
		REV	23	48	
PTEN_1	R130	FW	23	43	6
		REV	23	48	
SMAD2_1	S464*	FW	23	48	2
		REV	23	48	
SMAD3_1	R268H	FW	23	52	2
		REV	23	48	
SMAD4_1	R361, R356	FW	20	35	4
		REV	21	33	
SMAD4_2	D537	FW	20	50	2
		REV	20	55	
TGFBR2_1	K128, P129	FW	21	38	2
		REV	19	37	
TGFBR2_2	R528H	FW	22	50	3
		REV	23	48	
TP53_1	R273, R282	FW	21	52	6
		REV	19	58	
TP53_2	R248, G245	FW	20	55	6
		REV	19	63	
TP53_3	R175	FW	20	50	5
		REV	20	55	
TP53_4	R213, R196	FW	20	55	6
		REV	20	55	
TYMP_1	R81W	FW	20	55	3
		REV	20	55	
TYMS_1	G143R	FW	23	48	4
		REV	23	48	
UPP1_1	R297Q	FW	20	60	1
		REV	22	50	

2.4.3 Nested PCR protocol and library preparation

In order to reduce the DNA input, a custom nested PCR protocol suited for multiplex reactions was developed. For the first PCR step, at least 500pg of cirDNA were used per multiplexed reaction (if possible 1ng; for FFPE samples at least 1ng per reaction). In all cases Q5 Hot Start High-Fidelity Polymerase (New England Biolabs, Ipswich, USA) was used to minimize amplification errors. This polymerase has the highest fidelity so far, which is 280 times higher than Taq polymerase.[145] To reduce the formation of primer dimers a concentration of 40nM for each primer and a very long annealing time (6 min) were selected. This was in line with a previously published preamplification protocol for multiplex cirDNA-amplification.[146] PCR was performed using 5µl 5 x Q5 HS High-Fidelity buffer, 0.5 U Q5 HS High-Fidelity DNA Polymerase, 200µM dNTPS (Thermo Fisher Scientific), 40nM of each primer (Eurofins Genomics) and 0.5ng to 1.5ng DNA (for FFPE-samples up to 6ng DNA). The reaction volume was filled up to 25µl with nuclease-free water.

After an initial denaturation step (3 min, 98°C), 18 PCR cycles were run with the following settings: 10 s of denaturation at 98°C, 6 min of annealing at 62°C and 30 s of elongation at 72°C. In the end, a 10 min elongation step at 72°C was included. The PCR products were semi-automatically purified on a Biomek FX (Beckman Coulter, Brea, USA) using AmpureXP beads (Beckman Coulter) according to the manufacturer's instructions with a bead to sample ratio of 0.76 and eluted in 25µl of nuclease free water. By this step, fragments larger than 120bp were separated effectively from primer dimers.[147] In the second PCR step, the same settings were applied except for the following modifications: 15µl of DNA products from the first PCR step were taken as input amount. Primers consisting of Illumina universal adapter sequence and a unique combination of the TruSeq DNA HT indexes were used at a concentration of 300nM (respective primer sequences are listed in chapter 2.1.). This allowed sequencing of multiple samples (up to 96) per run. Third, the annealing temperature was increased to 72°C and the annealing time was reduced to 15 s.

After the second PCR step, all six multiplexed PCR reactions of each patient sample were pooled. 90µl of the mixed PCR products were then again semi-automatically cleaned with a Biomek FX using AmpureXP beads with an even lower bead to sample ratio of 0.67 and eluted in 20µl of nuclease free water. Thereby, fragments larger than 200bp were enriched.

Capillary gel electrophoresis with the Bioanalyzer 2100 (Agilent Technologies, Santa Clara, USA) was used to evaluate successful library amplification and library quality by assessing the DNA concentration of the relevant fragment sizes (235-310bp). Quality assessment yielded uniformly satisfying results, as far as derivable from the fragment size distribution. For a few samples, peaks at ~170bp indicated, that primer dimer products still were present, however, in comparably low amounts.

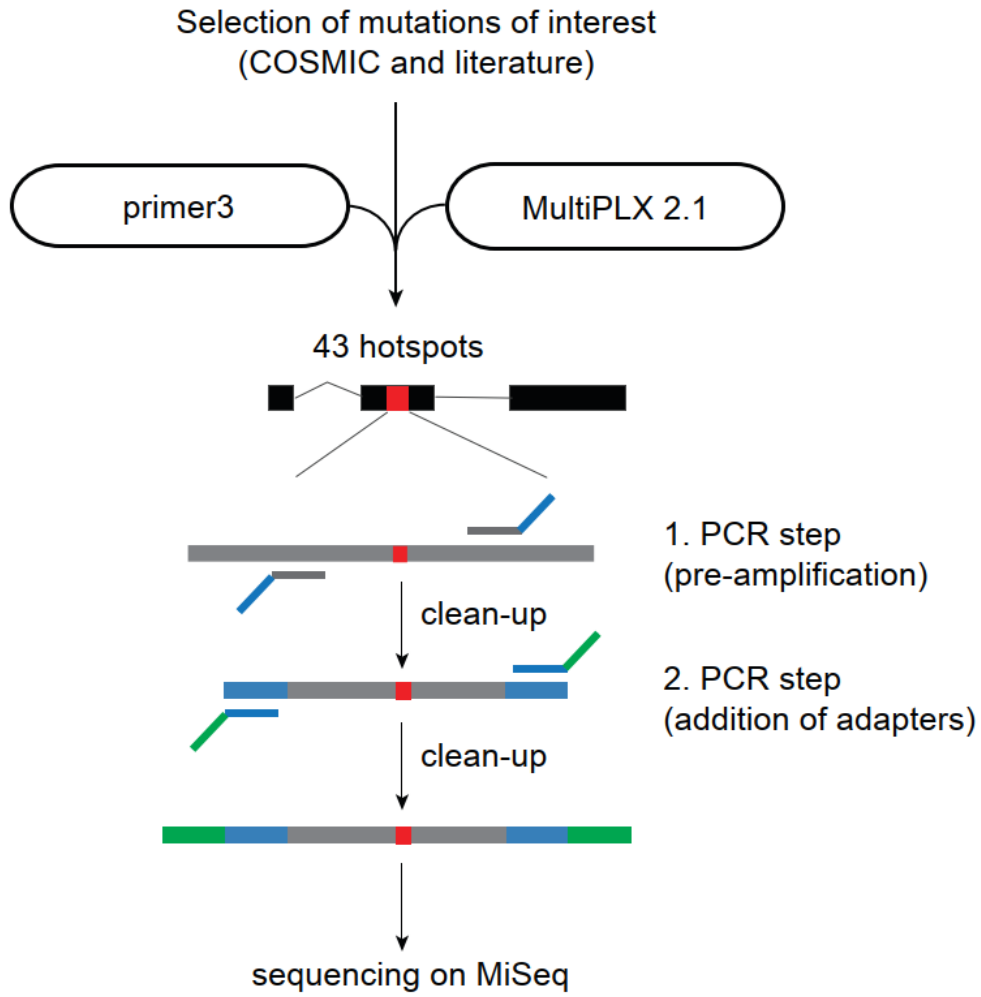

Figure 2-3. Experimental workflow of custom amplicon sequencing. Mutational hotspots of interest were selected in COSMIC database and cbioportal, amplicon primers were then designed with primer3 and multiplexes were determined by the MultipPLX 2.1 software. A two-step nested PCR workflow was established, including preamplification and addition of sequencing adapters before PCR clean-up and subsequent sequencing on MiSeq (Illumina).

2.4.4 Amplicon sequencing

All patient samples were pooled to a single library with a final concentration of 4nM. The library was sequenced on an Illumina MiSeq, using the MiSeq Reagent Kit V2 (150bp single end sequencing, Illumina, San Diego, USA) with 5% PhiX as spike-in. Between 30 and 50 patient samples were sequenced in one MiSeq run.

2.4.5 Analysis of sequencing data

Bioinformatic analysis was carried out by Benedikt Rauscher. Statistical analysis was performed in close cooperation. In the following a description as published in the manuscript is outlined.[148]

Quality assessment and variant calling

First, a quality report was generated for each sequenced sample using 'FASTQC'.[149] Quality reports were examined manually to ensure sufficient sample quality. Subsequently, the 'Trimmomatic' software version 0.36 was used to remove sequenced base pairs with a quality score of less than 15 (PHRED33) at both ends of the reads to avoid false positive mutations due to sequencing errors.[150] In addition, a sliding window trimming was performed, cutting the read once the average quality within a window of size 4bp was detected to be less than 15. Reads with a length of less than 70 base pairs were removed. Next, reads were mapped to the human genome GRCh38 using the 'bwa mem' algorithm of the 'bwa' alignment software version 0.7.15-r1140 with default parameters.[151] 'samtools' version 1.4 were used to remove reads that could not be mapped to genome.[152] In addition, reads mapped with a quality of less than 13 were excluded. To determine variants, the 'mpileup' algorithm implemented in the 'samtools' software was applied. The resulting variant calls were summarized and quantified using 'bam-readcount' (https://github.com/genome/bam-readcount), counting only variants at positions with a sequencing quality of at least 30.

Analysis of sequencing variants

The allele frequency was determined for each detected variant by dividing the number of reads containing the variant by the total number of reads mapping to that position. Subsequently, variants were mapped to the amplicon panel. To this end, the genomic coordinates were determined for each amplicon using the 'BLASTn' algorithm as implemented in the 'BLAST+' software package.[153] Here, each amplicon sequence was mapped to the human genome GRCh38 requiring 100% sequence identity. Sequencing variants were then matched to the custom amplicon panel by genomic coordinates. Correctness of the variant mapping and matching was assessed by taking advantage of prior knowledge about mutations in the *NRAS*, *KRAS* and *BRAF* oncogenes that had previously been determined using Sanger sequencing. In order to distinguish true mutations from false positive mutations introduced by, for example, PCR or sequencing errors, a model-based approach was applied. Assuming that a)

the majority of variants are caused by PCR or sequencing errors, that b) PCR errors do on average occur at the same frequency at each position of an amplicon and that c) sequencing errors at a specific genomic position are equally likely in each sample. A robust linear model of the form

$$\log\left(\frac{a}{1-a}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \epsilon$$

was fit for each amplicon where *a* is the allele frequency of a variant, X_1 is the variant's position on the amplicon and X_2 represents the sequenced patient sample. The resulting fit represents a noise model. A variant was considered a true mutation if its model residual exceeded the median of all residuals by at least three standard deviations indicating that its presence cannot be explained by the estimated noise present in the data. In addition, a minimum allele frequency of at least 0.5% for a variant to be considered a true clinically relevant mutation was required. Next, the remaining mutations were annotated using the COSMIC database of somatic mutations in cancer.[35] Mutations that were not listed in COSMIC were excluded from further analysis. In addition, COSMIC mutations marked as SNPs were excluded.

Estimate of ctDNA-levels

For subsequent analysis ctDNA-levels were assumed to be proportional to the maximal somatic MAF. Therefore, they were computed as mutated fraction of total cirDNA-level.

Statistics

All p-values reported in this thesis were computed using a two-sided student's t-test for parametric data and a two-sample Wilcoxon rank sum test or Kruskal-Wallis test combined with the uncorrected Dunn's test or Mann-Whitney test for non-parametric data. R statistical programming language or the GraphPad Prism software v. 8.1.0 were used. As metrics for the relationships between quantitative variables both the parametric Pearson correlation coefficient (PCC) and the rank-based Spearman correlation coefficient (SCC) were computed. Venn diagrams were made using BioVenn.[154]

Data and software availability

All sequencing data generated are available from the European Genome-phenome Archive under the accession number EGAS00001003382. Documented computer code to reproduce sequencing analyses presented is available from GitHub at

https://github.com/boutroslab/Supplemental-Material/tree/master/Herrmann%26Zhan%26Betge%26Rauscher_2018.

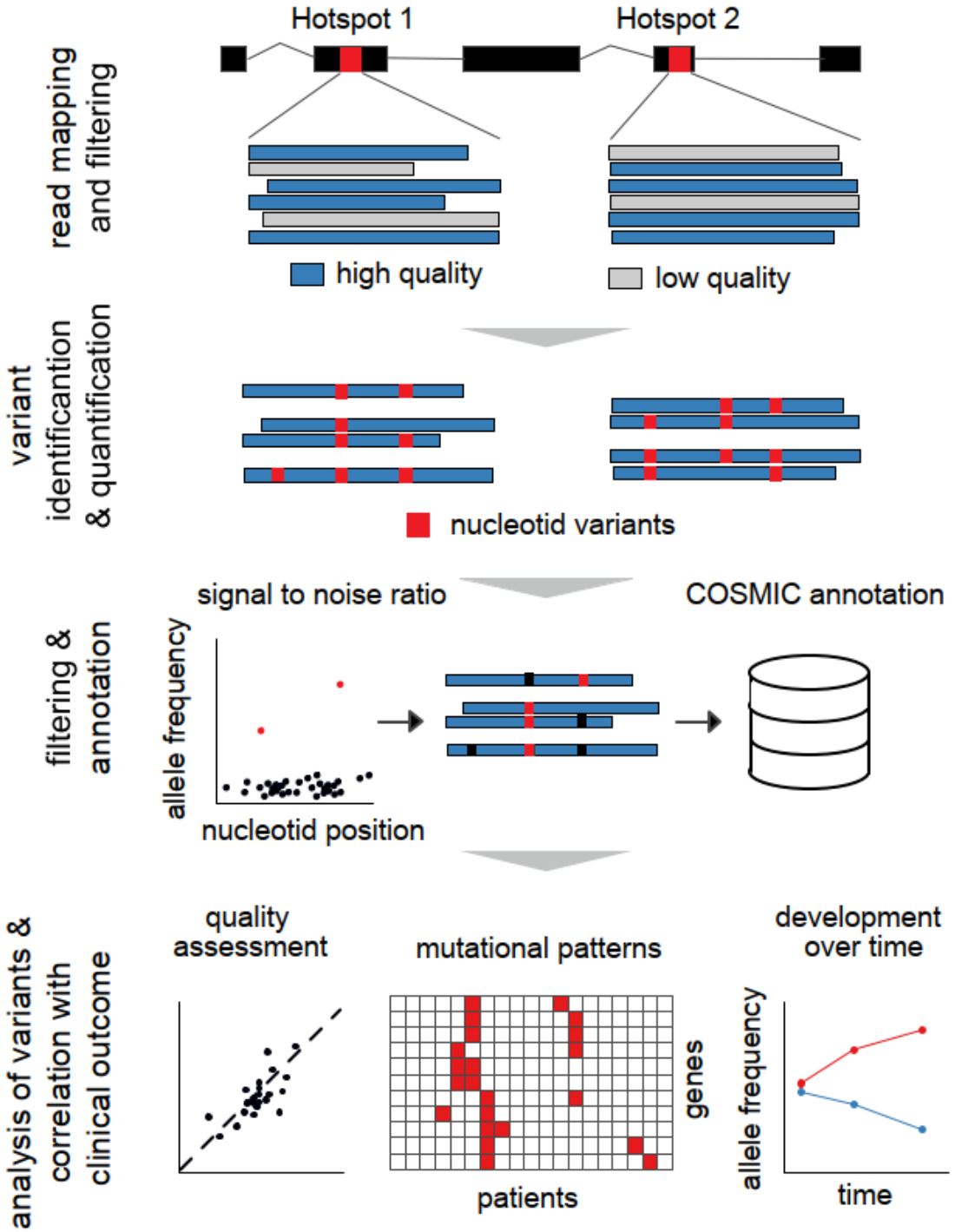

Figure 2-4. Bioinformatic analysis workflow. Reads generated with Illumina MiSeq were mapped to the human reference genome and low-quality reads were excluded before variant identification by comparison of reads and reference. Allele frequency was calculated, and variants were filtered by signal to noise analysis. Subsequently, variants were annotated with metadata from public databases. Importantly, variants not present in the COSMIC database were filtered out. With this set of mutations, further analyses regarding mutational patterns or the development of mutations over time were performed.

3 RESULTS

3.1 Circulating DNA as a tumor biomarker.

3.1.1 Composition of the MALIBU study cohort

To study the value of cirDNA as a biomarker in gastrointestinal cancers, serial blood samples from cancer patients were collected prospectively at the University Hospital Mannheim, Heidelberg University within the Mannheim Liquid Biopsy Unit (MALIBU) (Figure 2-1). For patients undergoing palliative chemotherapy, blood samples prior to the start of treatment and in parallel to each radiological assessment of therapy response were obtained. In all cases blood sampling was performed before administration of anti-cancer drugs. For patients undergoing surgical removal of the cancer, blood samples were collected prior and at multiple time points after resection. In parallel, relevant clinical parameters and blood markers were documented in a prospective database. In total, blood samples from 280 patients with various GI-tumors across all UICC stages were collected, predominantly from patients with metastatic disease (Figure 3-1A-C). Colorectal cancer represented the most common tumor entity with 37% and a substantial part of this cohort had advanced stages of disease with multiple distant metastases (Figure 3-1A, F-G).

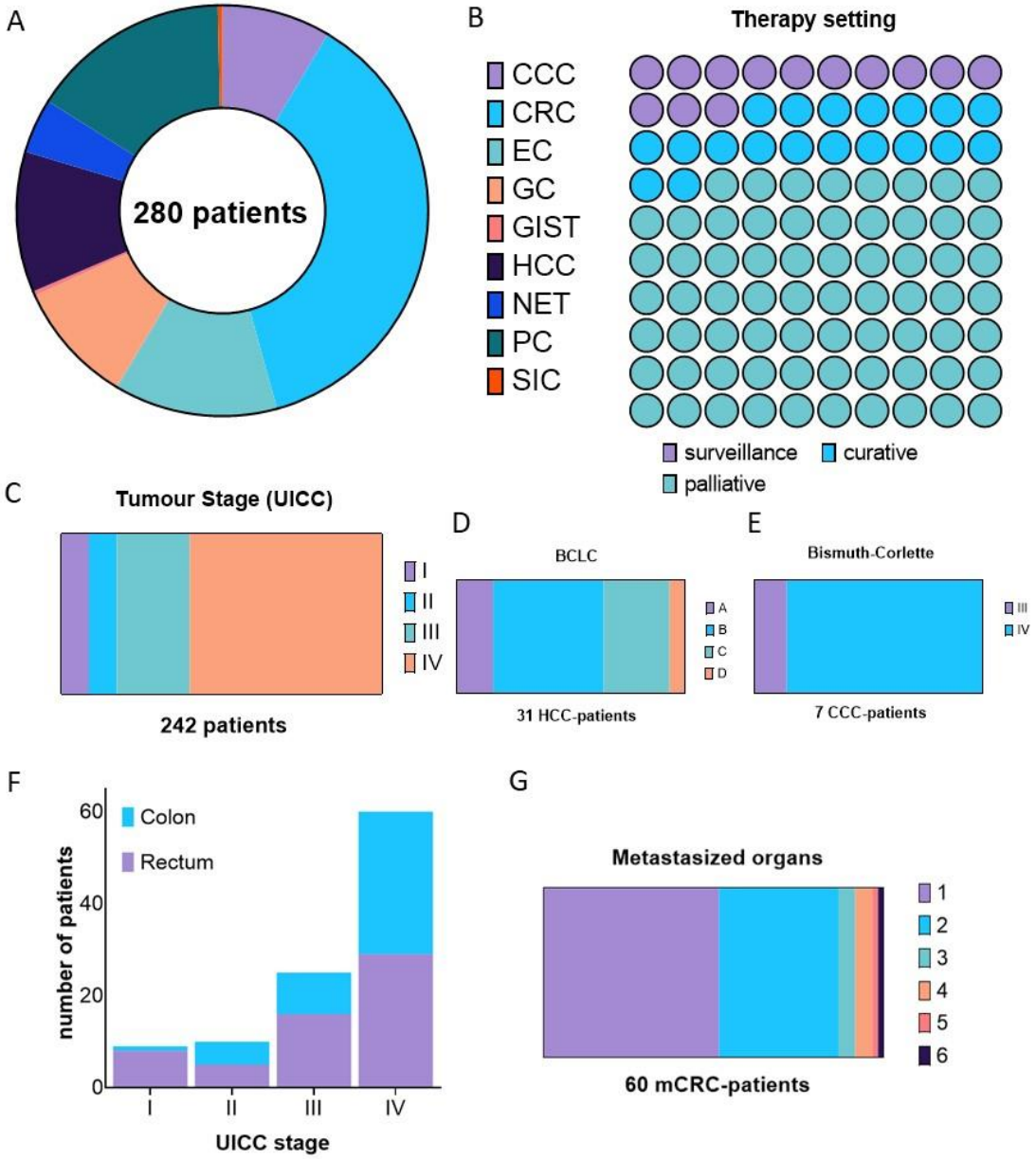

Figure 3-1. Characteristics of patients within the MALIBU study cohort.
(A) Tumor entities within the MALIBU cohort. The cohort comprises patients with cholangiocellular carcinoma (CCC), colorectal cancer (CRC), esophageal cancer (EC), gastric cancer (GC), gastrointestinal stroma tumor (GIST), hepatocellular carcinoma (HCC), neuroendocrine tumors of the GI-tract (NET), pancreatic cancer (PC) and small intestinal cancer (SIC).
(B) Therapy of patients within the MALIBU cohort. About two-thirds of patients received palliative treatment and care.
(C) Staging according to the Union for International Cancer Control (UICC) classification. Patients with early and late-stage diseases were included.
(D) Staging of HCC-patients according to Barcelona Clinic Liver Cancer (BCLC) classification.
(E) Classification of patients with hilar cholangiocarcinoma (Klatskin-tumors) according to Bismuth-Corlette.

(F) Stage and location of patients with CRC. More than half of the patients showed distant metastases (stage IV) by the time of inclusion.
(G) Half of CRC patients had metastases in two or more organs.

3.1.2 Fragment size of cirDNA in gastrointestinal cancer patients

To determine the constitution of cirDNA fragmentation analysis was performed on selected patient samples. Almost all samples exhibited a prominent peak at ~166bp, representing the fragment size of a chromatosome unit. Also, some cirDNA samples showed a substantial amount of long DNA fragments of 2000bp up to 10000bp.

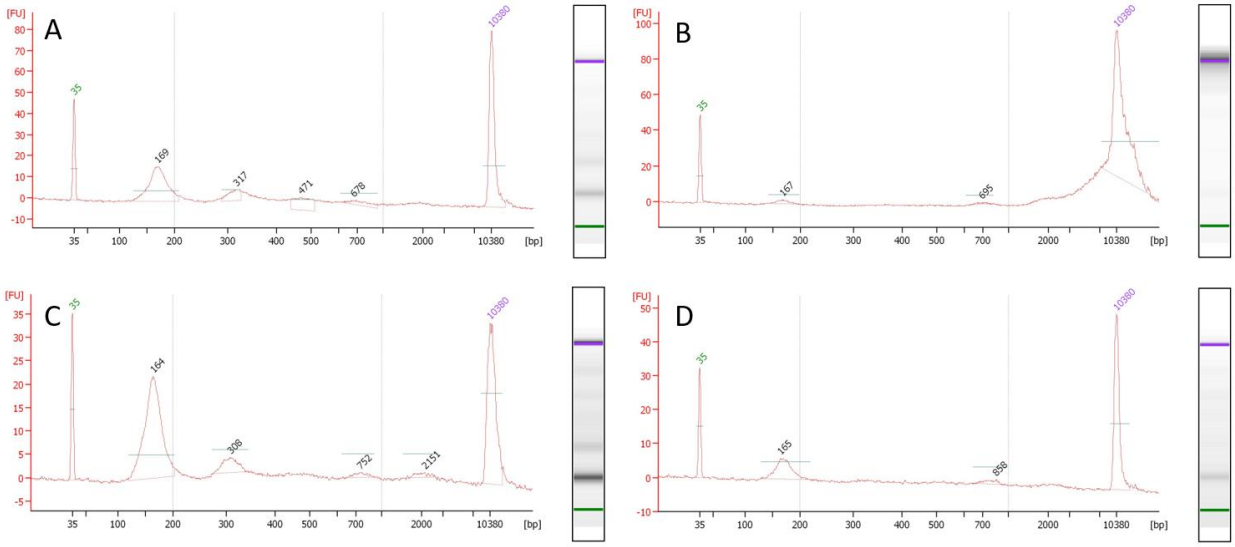

Figure 3-2. Examples of cirDNA fragment size.
(A) Chromatosome sized peaks can be identified with the singular unit (169bp) being the dominant fraction. This patient had a very high concentration of cirDNA (177.02ng/ml plasma) and a large proportion of ctDNA (91% of all reads carried the truncating mutation R213* in TP53).
(B) Dominant fractions were large fragments >2000bp, probably deriving from extensive cell necrosis. The cirDNA-level was 57.07ng/ml plasma.
(C) Fragments lengths resemble the chromatosome units. cirDNA-level was 9.60ng/ml plasma.
(D) Fragments about 165bp long are identifiable. cirDNA-level was 2.82ng/ml plasma.

3.1.3 cirDNA-levels in different gastrointestinal tumors

To analyze associations of cirDNA amounts with clinical parameters, cirDNA concentrations from all study patients were measured. A total of 389 cirDNA samples of 280 patients were analyzed.

No differences in cirDNA-levels were observed between sexes and tumor entities (Figure 3-3A-B). Advanced tumor stages were associated with elevated cirDNA-levels. This was also observed when comparing the cirDNA amounts of patients in different

settings, as palliatively treated patients showed higher levels of cirDNA (Figure 3-3C-D). Tumor free patients had the lowest cirDNA-levels, but also patients with a therapy response in radiologic imaging showed lower levels as compared to those who had progressive disease (Figure 3-3E).

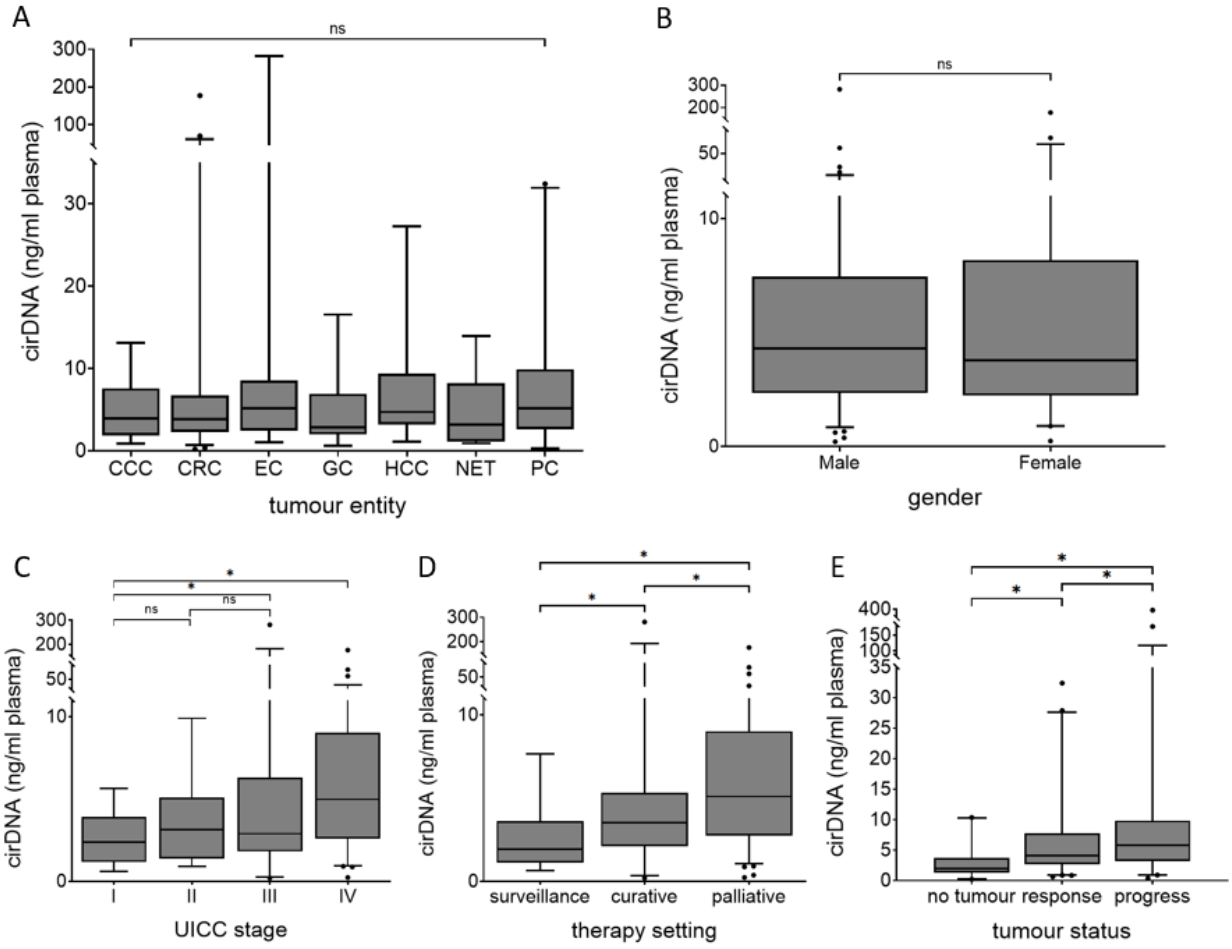

Figure 3-3. cirDNA-levels in the MALIBU cohort. The first sample of each patient was considered. 97.5 percentiles are depicted.
(A) cirDNA-levels in patients with different gastrointestinal tumors are displayed. No difference was also observed between entities correcting for tumor stage (not shown).
(B) cirDNA-levels of female and male patients.
(C) UICC stage and corresponding cirDNA-levels. Patients with HCC and Klatskin-tumors were excluded.
(D) cirDNA-levels by therapy setting, categorized as curative therapy (neoadjuvant and adjuvant treatment), palliative therapy and surveillance after curative treatment.
(E) cirDNA-levels and tumor status as assessed by imaging studies. Every sample with an available radiological staging was included. Response is defined as either regressive or stable disease.

Analyzing only the subgroup of CRC patients, again patients in stage IV showed higher levels of cirDNA as well as patients treated palliatively (Figure 3-4).

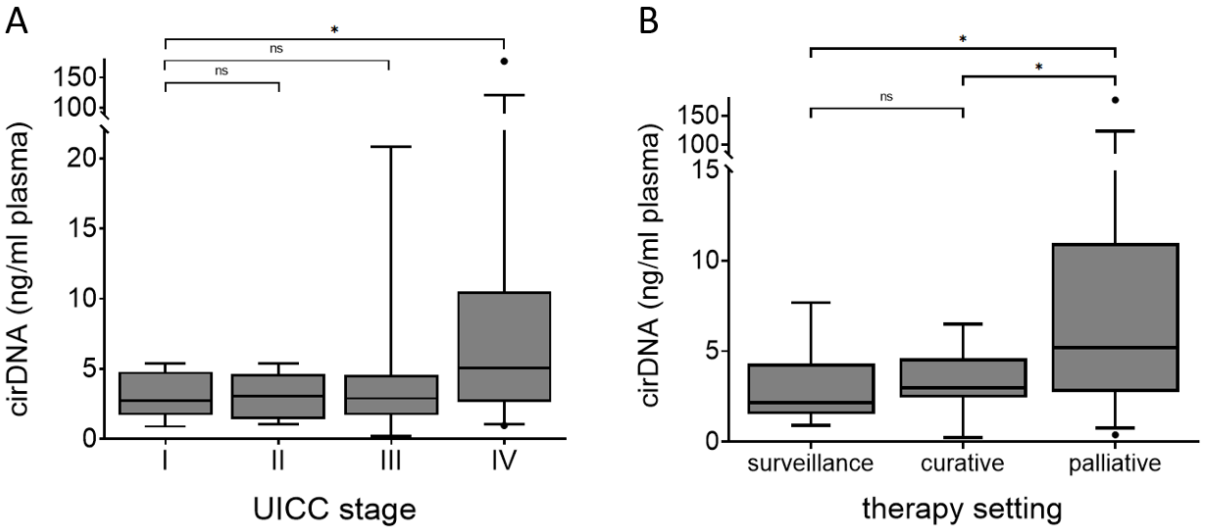

Figure 3-4. cirDNA-levels in the CRC-cohort.
(A) cirDNA-levels grouped by UICC stage of the patient's tumor.
(B) cirDNA-levels grouped by the treatment intent.

cirDNA-levels showed multiple associations with blood parameters. Higher cirDNA-levels correlated with higher LDH values (Figure 3-5A). However, this correlation did only account for samples with LDH levels >500U/l (PCC=0.870, p=0.0005; SCC=0.791, p=0.0055; n=11). For samples with a LDH <500U/l no Pearson correlation was evident (PCC=0.005, p=0.932; SCC=0.235, p=0.0002). When LDH-levels were differentiated as elevated vs. not elevated with a default cut-off of 300U/l cirDNA-levels differed with a median of 4.21ng/ml plasma for non-elevated LDH and 6.04ng/ml plasma for elevated LDH values (p=0.001). Likewise, CRP values and cirDNA-levels showed a positive correlation (Figure 3-5B). These observations held true for subgroup analyses of CRC patients (Figure 3-6A-B). Weak correlation was noticed between CEA and cirDNA-levels. This was stronger when only CRC samples were analyzed (Figure 3-5D and Figure 3-6D). CEA is the established tumor marker in CRC and has limited use in patients with other tumors. The same applied to albumin levels. Here lower levels correlated with rising cirDNA-levels (Figure 3-5C and Figure 3-6C).

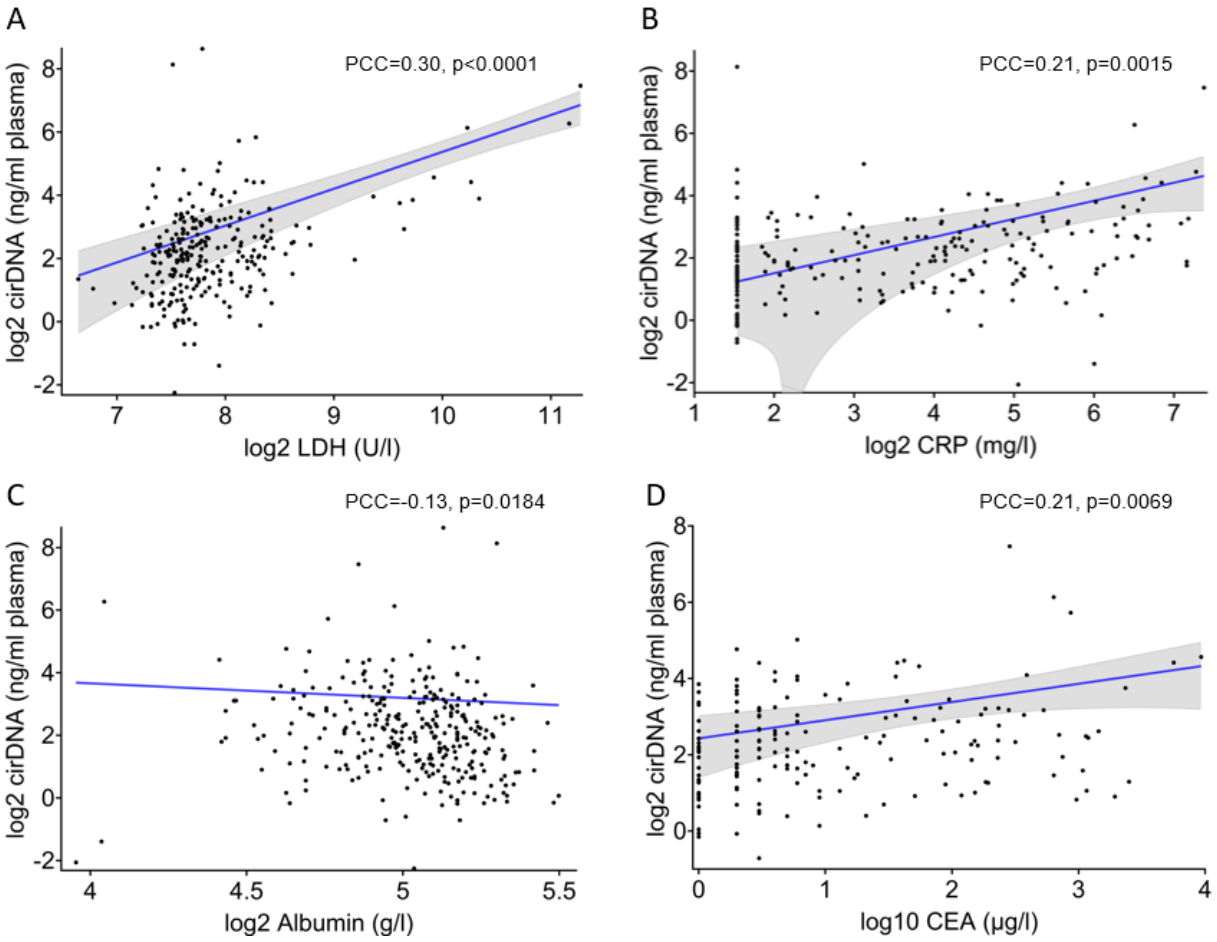

Figure 3-5. Correlations of blood parameters with cirDNA-levels in patients with gastrointestinal tumors. Logarithmic axes were chosen for easier visualization. 95% confidence bands are sketched as grey areas.
(A) Association of cirDNA-levels with LDH-levels. This correlation was, however, limited to a few outliers. For most of the LDH values no correlation was seen.
(B) Association of cirDNA-levels with CRP values. The assay for CRP measurement had a lower detection limit of 2.9mg/l. This limit was taken as default value for samples with no detectable CRP.
(C) Association of cirDNA-levels with albumin. A weak negative correlation was observed. Here, no confidence bands are shown due to the very weak correlation.
(D) Association of cirDNA-levels with serum tumor marker CEA.

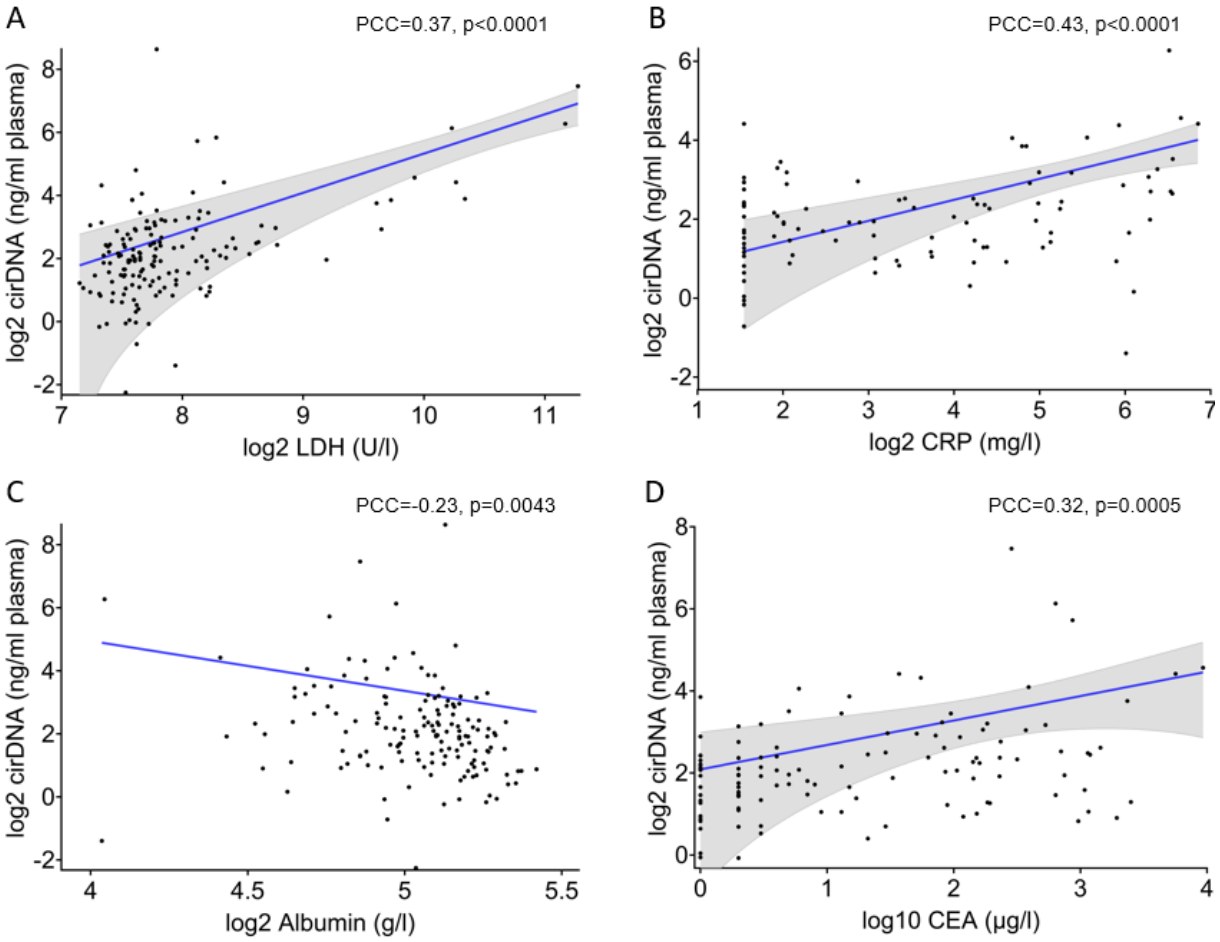

Figure 3-6. Correlations of blood parameters with cirDNA-levels in CRC patients. Logarithmic axes were chosen for easier visualization. 95% confidence bands are sketched as grey areas.
(A) Association of cirDNA-levels with LDH-levels. This correlation was, however, limited to a few outliers. For most of the LDH values no correlation was seen.
(B) Association of cirDNA-levels with CRP values. The assay for CRP measurement had a lower detection limit of 2.9mg/l. This limit was taken as default value for samples with no detectable CRP.
(C) Association of cirDNA-levels with albumin. A weak negative correlation was observed. Here, no confidence bands are shown, due to the very weak correlation.
(D) Association of cirDNA-levels with serum tumor marker CEA.

Overall correlations between cirDNA and other blood parameters were weak. No other relevant correlation was observed considering age, other laboratory parameters (Hb, WBC, platelets, INR, creatinine and bilirubin) and tumor markers AFP and CA19-9.

3.2 Establishing amplicon sequencing for cirDNA samples and application to a cohort of patients with late-stage colorectal cancer.

3.2.1 Development of a multiplex amplicon-based PCR protocol

To analyze mutational patterns in cirDNA, a pipeline for the design of custom amplicon panels was developed. Primers that cover mutational hotspots in oncogenes and tumor suppressors of CRC were selected based on the COSMIC and cbioportal database.[33,35] In addition, loci in genes that affect metabolism of 5-FU[70,71,73] or DNA mismatch repair were included.[26] These mutations occur at a lower frequency and are not commonly covered by commercial amplicon panels. An overview of all genes can be found in Figure 2-2. Then, primer pairs that bind ~50bp up- and downstream of the mutations of interest were designed and optimal multiplexes calculated (see Methods for details). The custom panel contained a total of 43 primer pairs, which were distributed over six multiplex reactions containing 5-9 primer pairs each (see Table 2-4 and Chapter 2.1 for primer characteristics and associated hotspot mutations).

Primer validation

Primers were designed with uniform parameters to ensure similar performance under PCR conditions. Emphasis was put on a small range in melting temperatures (55°C to 61.5°C).

First all designed primers were tested in a single-plex PCR to verify correct target amplification. This was performed using both genomic DNA and cirDNA. By agarose gel electrophoresis DNA fragment size was determined. Primers not showing right sized products were discarded and a new primer set was designed. Also, primers predominantly amplifying dimers seen as small sized fragments (~100bp) in gel electrophoresis were dismissed.

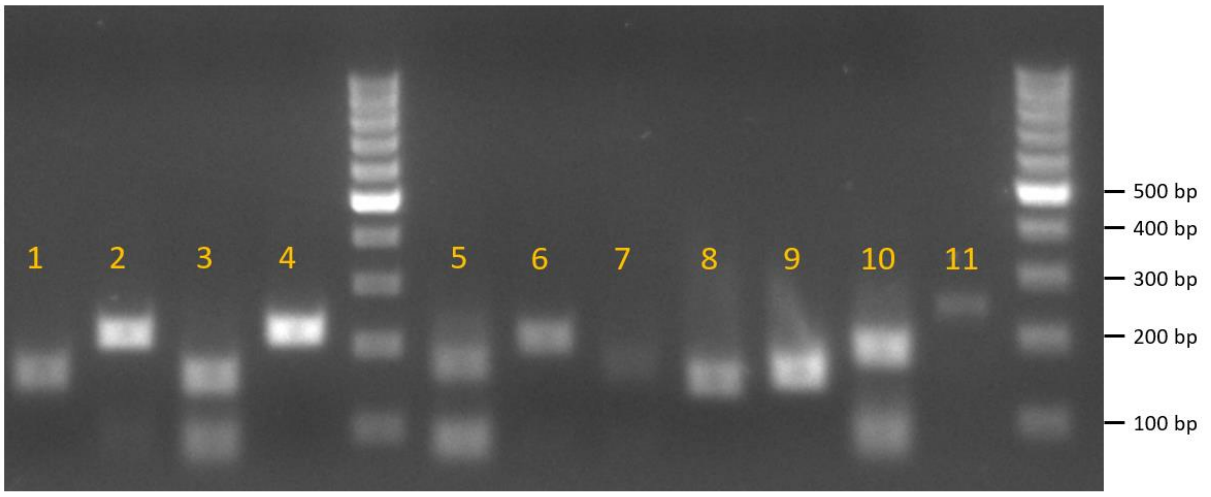

Figure 3-7. Results of an agarose gel electrophoresis (exemplary). Samples were loaded into the pockets after the first PCR amplification. Expected amplicon sizes are between 150 and 220bp. *1-2* both targeted FBXW7_1. *3-4* targeted TGFBR2_1. *5-6* targeted SMAD4_1. *7-8* targeted SMAD4_2. *9* amplifies ERBB3 and *10* PIK3CA_3. *11* is a control with 270bp fragment length. Given these results *1,3,5,7* were dismissed due to inferior performance.

Multiplex-PCR optimization and functional validation

Primer dimer formation and amplification in multiplex assays was minimized by drastically reducing the amounts of all primers down to 40nM each (~10 times lower than in a standard assay). Under these conditions, primers needed more time to locate their respective target. Thus, the annealing time was lengthened to 6 min. This was previously described as a preamplification step.[146]

Finally, the overall number of cycles was reduced to 36. This still yielded enough products, minimized the amount of PCR errors and thereby simplified the sequencing data analysis. By cutting down the cycle count of the first PCR from initially 25 cycles down to 18, discrepant amplification efficiencies were given less influence. Thereby, the variance between amplicons could be kept in tolerable ranges (sequencing depth of amplicons; minimum 503, maximum 20809, median 7576).

PCR products were semi-automatically purified on a Biomek FX using Ampure XP beads. By changing the sample-to-bead ratio, purification could be focused on specific size ranges. Different ratios for each clean-up step were tested considering the respective amplicon length (150-225bp after preamplification and 235-310bp after second PCR). By applying 0.76µl beads per 1µl PCR-product unwanted products of ~100bp length could be washed out by >92% (relative reduction >1300%) as determined by the Bioanalyzer 2100 (Figure 3-8A-B). A ratio of 0.66 was used for the

second clean-up. Again, correct amplification products could be enriched (reduction of ~170bp fragments by at least 45% (relative reduction >180%) (Figure 3-8C-D). Elution volume was reduced (to 25µl after first PCR and 20µl after second PCR) to have higher concentrations of substrates for following steps. Multiple samples were pooled to create a library ready to sequence. A last quality check showed that amplicon libraries had a purity of 82-90% (Figure 3-8E).

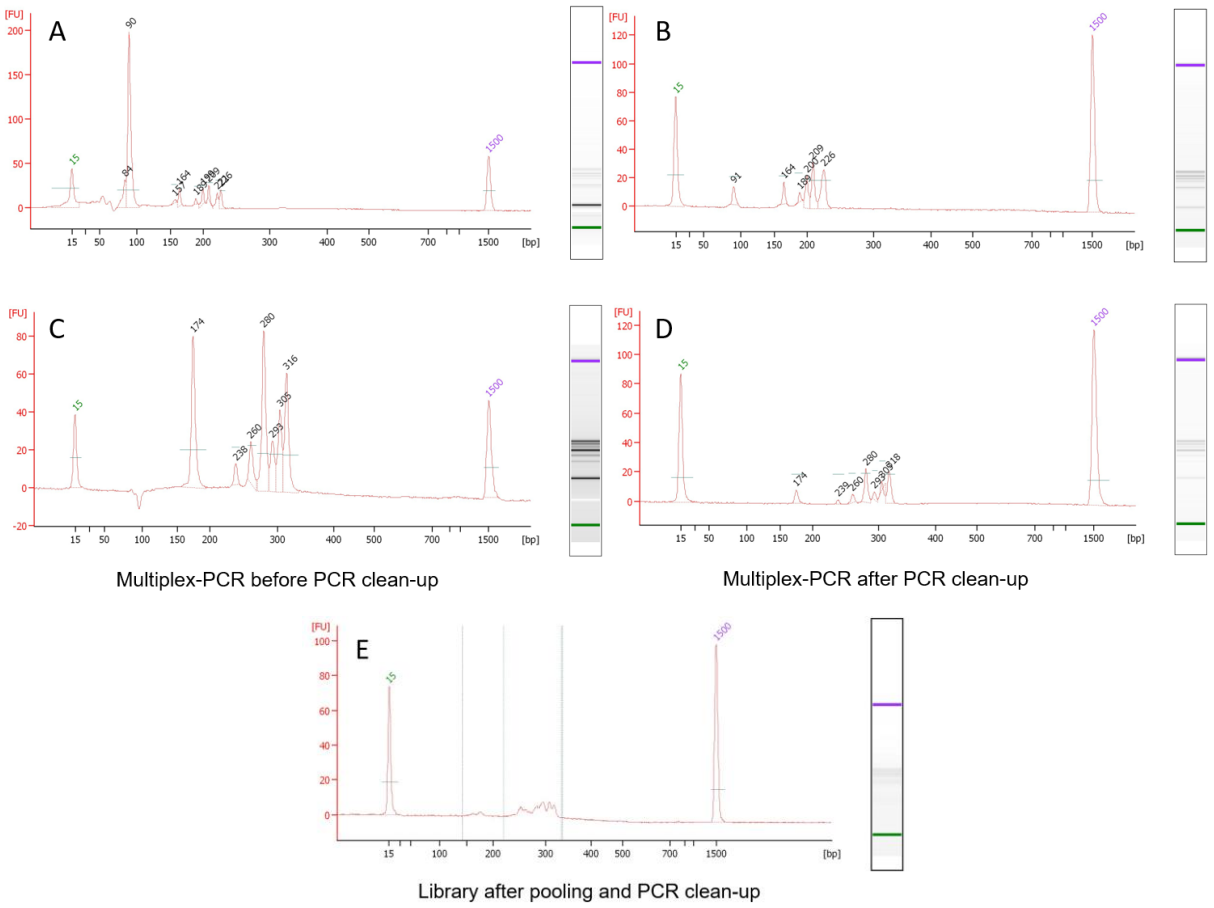

Figure 3-8. Efficiency of PCR purification assays.
(A-B) Example of pre- and post-purification results after the first PCR.
(C-D) Example of pre- and post-purification results after the second PCR.
Formation of dimers varied in between samples. Here two samples with high numbers of false products are shown, in which reduction of primer dimers could be achieved.
(E) A sequencing library consisting of 40 patient samples. Here, size regions are marked, and DNA concentrations are calculated by the Bioanalyzer software. Target regions make up for >82% of library DNA.

In order to find a multiplex concept for the panel the best fit model was predicted with the MultiPLX 2.1 software.[144] Further testing with varying input amounts of genomic and cirDNA was done and multiplex libraries were then analyzed by capillary gel electrophoresis and subsequent sequencing on an Illumina MiSeq. Figure 3-9 shows

fragmentation analysis of the six different multiplex reactions after the second clean-up step. In the standardized workflow the multiplex reactions were pooled before the second purification step in order to reduce workload and costs per sample (see Methods).

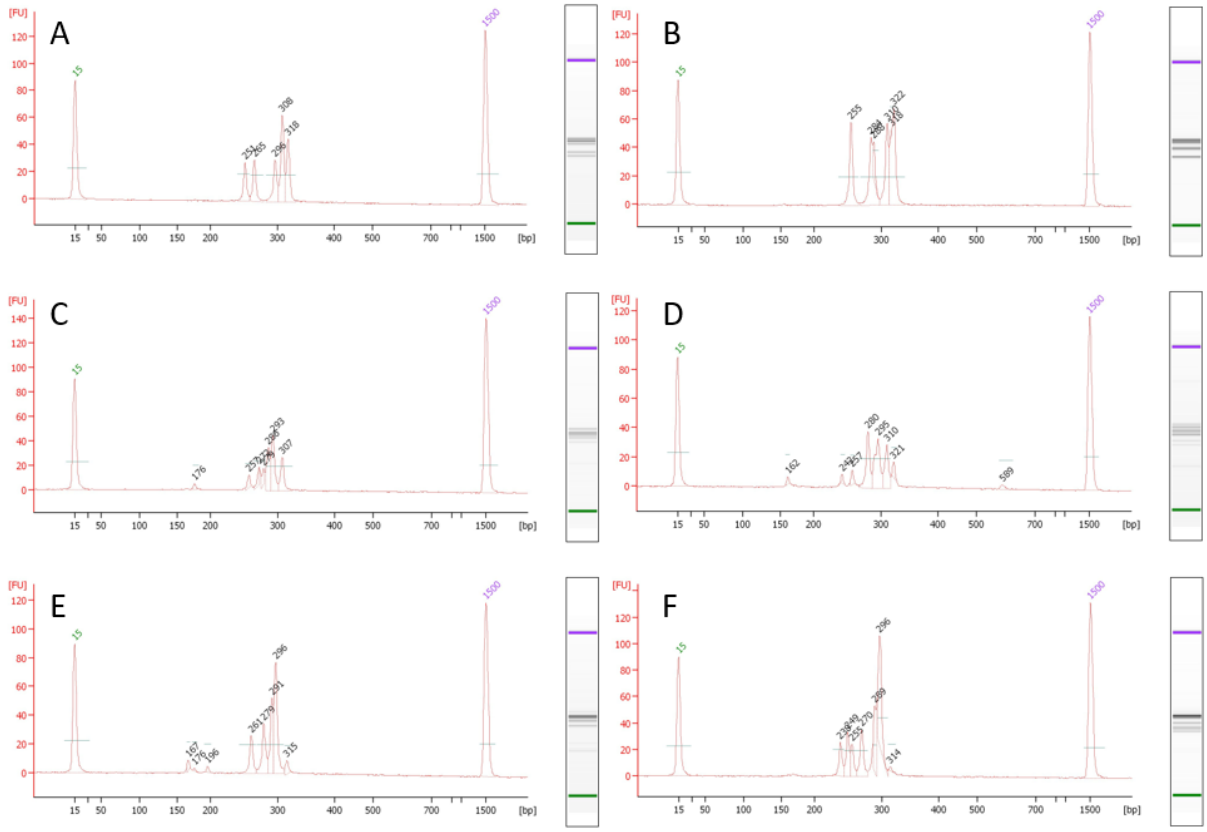

Figure 3-9. Fragment size distribution of the six multiplex reactions (**A-F**). The respective peaks of the amplicons can be distinguished as far as they differ in length, giving an appropriate impression of the amplification success. Formation and amplification of primer dimers could not be avoided completely, however, their concentration was comparably low. Overall, DNA concentrations between the assays varied only to a moderate extent, so that pooling of the multiplex mixes could subsequently be conducted blindly.

To validate the sensitivity of the method two different colorectal cancer cell lines with known mutations covered by the panel were mixed. Mutations could reliably be detected down to 0.62% allele frequency (Figure 3-10A). Therefore, the detection threshold for mutations was set at 0.5% MAF. The multiplex assay showed overall good performance both with 500pg (replicate 1&3) and 1000pg (replicate 2) of template DNA. In total only 3ng of cirDNA was needed to perform the multiplex assay. Hence, even patients with extremely low amounts of cirDNA could be included.

The comparison of FFPE samples with Sanger sequencing further validated the approach and pointed out that the method also worked on DNA extracted from archived histological material. All six known mutations in KRAS and NRAS covered by the panel

could be detected. Patients 133 and 167 had a KRAS A146T mutation in Exon 4. Only amplicons for the hotspots in Exon 2 and 3 were designed and thus this mutation could not be discovered (Figure 3-10B).

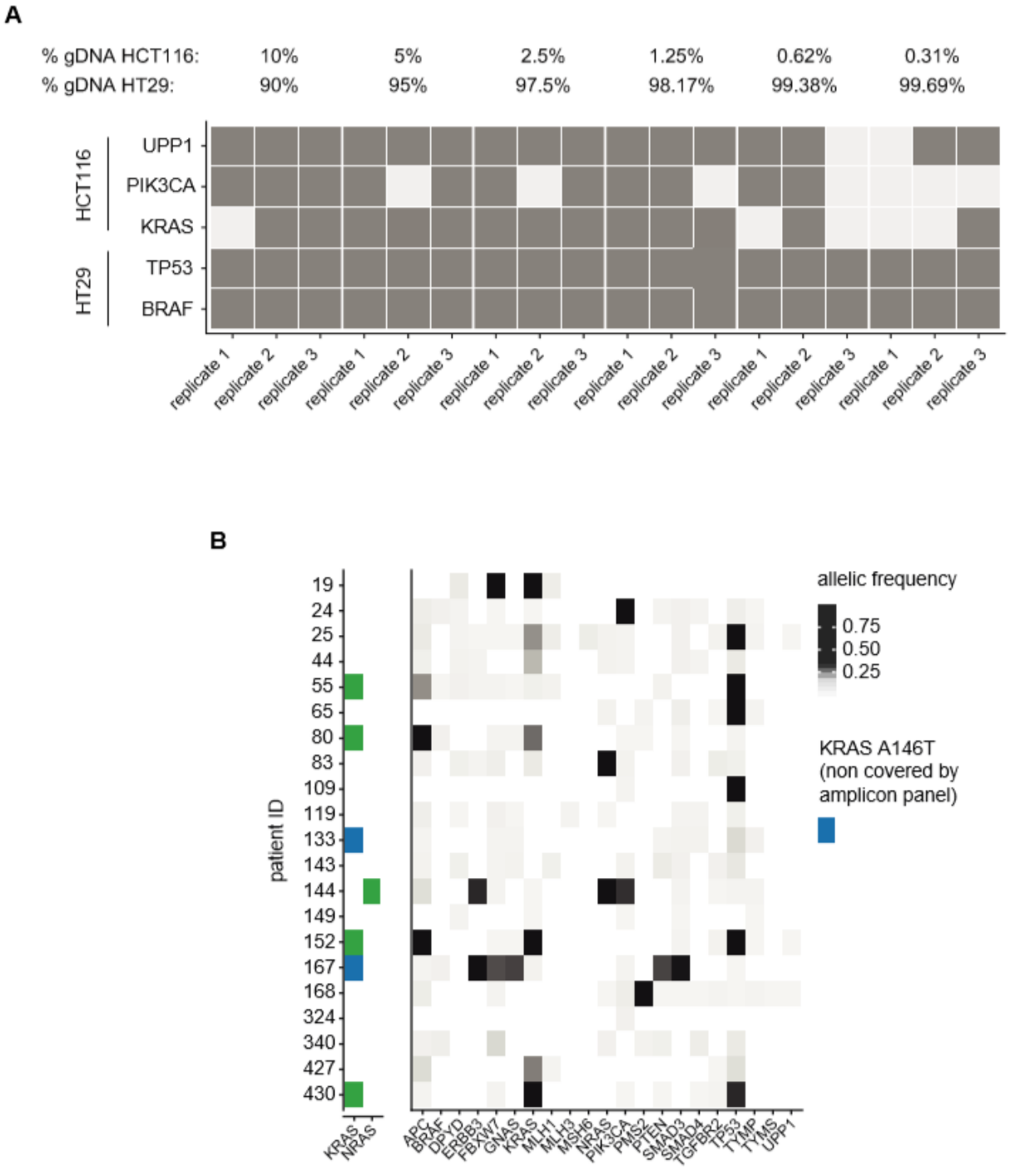

Figure 3-10. Sensitivity of the amplicon sequencing assay.
(**A**) DNA isolated from HCT116 cells with known KRAS, PIK3CA and UPP1 mutations were mixed with HT29-DNA (with known BRAF and TP53 mutations) in decreasing concentrations. The sequencing assay securely detected mutations down to a fraction of 1.25% HCT116 DNA (i.e. an allele frequency of 0.62%). MSH6 mutation was excluded due to dysfunctional amplification.
(**B**) Analysis of primary tumor samples. Amplicon sequencing (right) was performed analogous to liquid biopsy samples on DNA isolated from FFPE-tissue. Sanger sequencing (left) for

hotspot mutations in KRAS and NRAS had been routinely performed with DNA isolated from FFPE for therapy planning.

3.2.2 A pipeline for the analysis of amplicon-based sequencing of cirDNA

A custom bioinformatics pipeline for the analysis of sequencing results was developed (Figure 2-4). This pipeline comprised four consecutive steps. First, quality control in which reads with low sequencing quality were discarded. In the second step, all nucleotide variants across all reads of the same amplicon were detected and quantified. To discriminate PCR or sequencing artefacts from true genetic variants, a robust linear regression model was applied to each individual amplicon. Alternations identified as true variants by the model were then matched to the COSMIC database and discarded if not covered. By this bioinformatic approach, it was assured that only highly confident alterations previously found in cancer tissue were reported.

Mapping of sequencing reads to the amplicons showed that the amplicon coverage was continuously distributed (median 7576, minimum 503, maximum 20809). The extremes varied 41-fold (Figure 3-11A). The number of mutations identified varied between the different samples (Figure 3-11B). Also, a correlation between sequenced reads and the number of mutations per amplicon was noticed (Figure 3-11C).
The reproducibility of the experimental and analysis pipeline was validated by performing amplicon sequencing on independent biological replicates of 19 cirDNA samples, starting from cirDNA isolation from plasma. Mutations that occurred at an allele frequency >0.5% could reliably be detected, as indicated by a strong correlation between allele frequencies of the same mutations from independent replicates (Figure 3-11D). Nevertheless, dropouts were noticed. A few low frequent mutations were not present in the corresponding replicate. Failure of the assay was mainly restricted to a small set of amplicons or individual samples (Figure 3-11E-F).

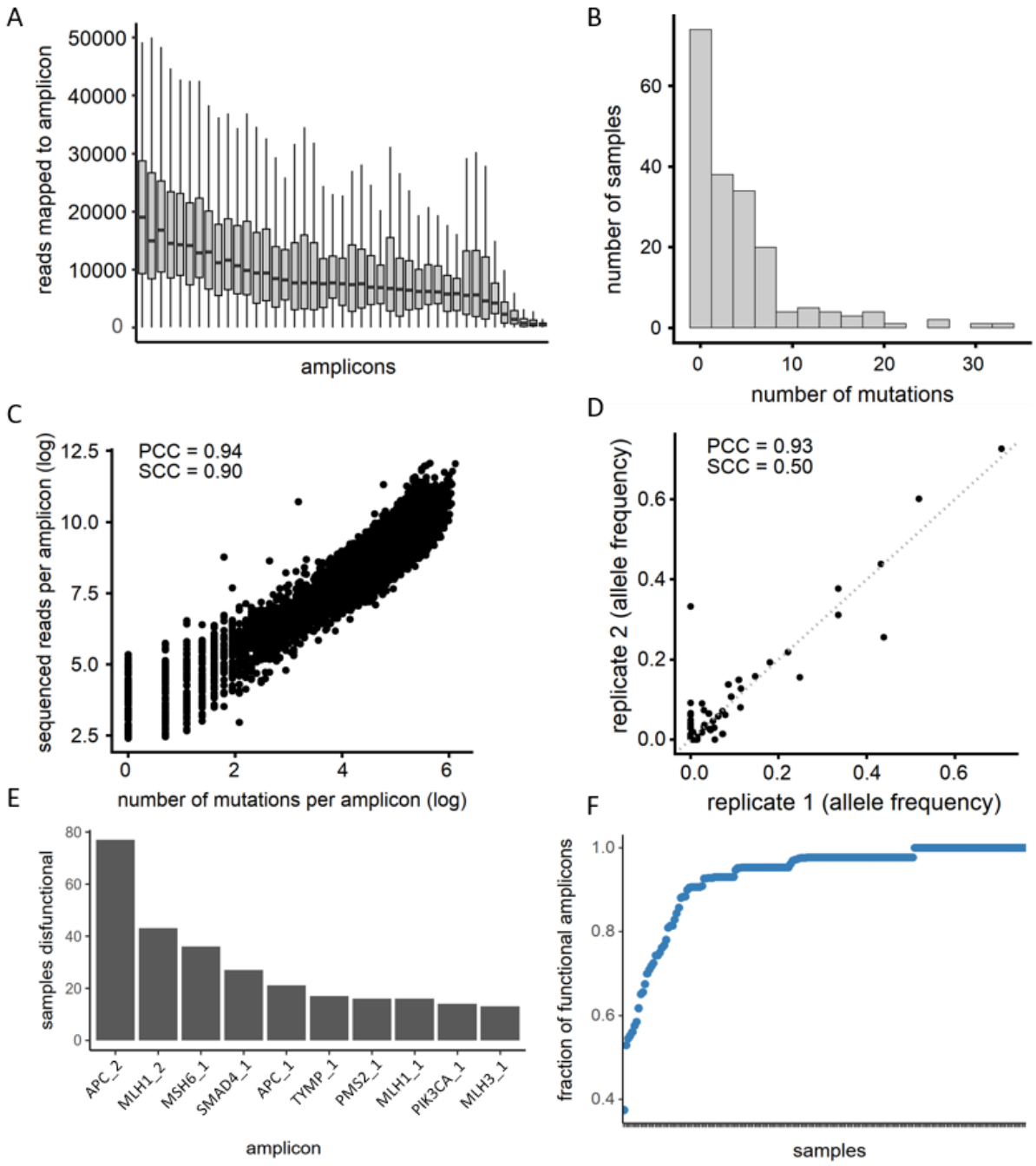

Figure 3-11. Performance indicators of multiplex assay and bioinformatic pipeline.
(A) Distribution of reads on amplicons. The number of reads mapped to each amplicon region is shown for all 43 amplicons. Data from all cirDNA samples was used for this analysis.
(B) Number of mutations per sample. Only mutations with positive signal-to-noise ratio and present in COSMIC database were considered true somatic mutations.
(C) Association between number of mutations found per amplicon and sequencing depth. Raw mutations were used for this analysis without any filter steps.
(D) Correlation of replicates. Two independent replicates were analyzed, starting from cirDNA isolation from the plasma. Shown are the allele frequencies of detected mutations in replicate 1 vs. replicate 2 of 19 analyzed patient samples.
(E) Most dysfunctional amplicons.
(F) Individual sample performance. Few samples displayed low fractions of functional amplicons.

In summary, quality control experiments indicated that custom amplicon sequencing was functional with cirDNA and genomic, including DNA isolated from archived FFPE samples. Furthermore, thresholds for MAF that enable the detection of genetic alterations with high confidence were identified.

3.2.3 Comparison of liquid biopsy and tissue biopsy in the mCRC-cohort

Amplicon sequencing of tumor tissue enables superior characterization

Mutations called by Sanger sequencing could all be detected by amplicon sequencing. (Figure 3-12A) Moreover, 31 mutations in RAS and BRAF were found by amplicon sequencing only, underlining the higher sensitivity of the next-generation sequencing technique. Sanger sequencing of BRAF was only performed in 10 cases, therefore four BRAF mutations found by amplicon sequencing, might have also been detected by Sanger sequencing.

Liquid biopsy detects mutations years after their first description

Known mutations in the MAPK signaling pathway, as determined by routine Sanger sequencing of tumor tissue, were retrieved in liquid biopsy samples in 9/16 cases (6/12 KRAS G12 or G13, 1/2 NRAS Q61, 2/2 BRAF V600E) (Figure 3-12B). Histological samples were taken 1.77 years before the first liquid biopsy in the mismatched group and 1.24 years for those that were rediscovered (matched group). The difference was not statistically significant (p=0.44). cirDNA-levels differed between both groups (p=0.0162). Samples in which RAS mutations could not be detected also had lower cirDNA concentrations (median 3.79ng/ml vs 5.24ng/ml plasma).

Interestingly, the MAF between the mismatched and matched group varied tremendously. The mean MAF was 14.73% in the matched group of MAPK-mutations compared to only 1.17% in the mismatched group. In 7/8 patients, recurrent mutations were found in the matched group (between 1 and 4; p=0.04). By contrast only two patients in the negative group exhibited a recurrent mutation (16 vs. 2 recurrent mutations; p=0.01). Patient 24 was not considered as she exhibited one recurrent mutation found in her colon primary tumor, but the KRAS mutation of the lung cancer was not detected, thereby assuming only cirDNA of the CRC was present.

17 additional mutations in KRAS, NRAS and BRAF were found in cirDNA samples, that were not detected in tumor tissue by Sanger sequencing (Figure 3-12B). These had considerably lower frequencies than matched MAPK-pathway mutations (median 1.20 vs. 8.60; p=0.009).

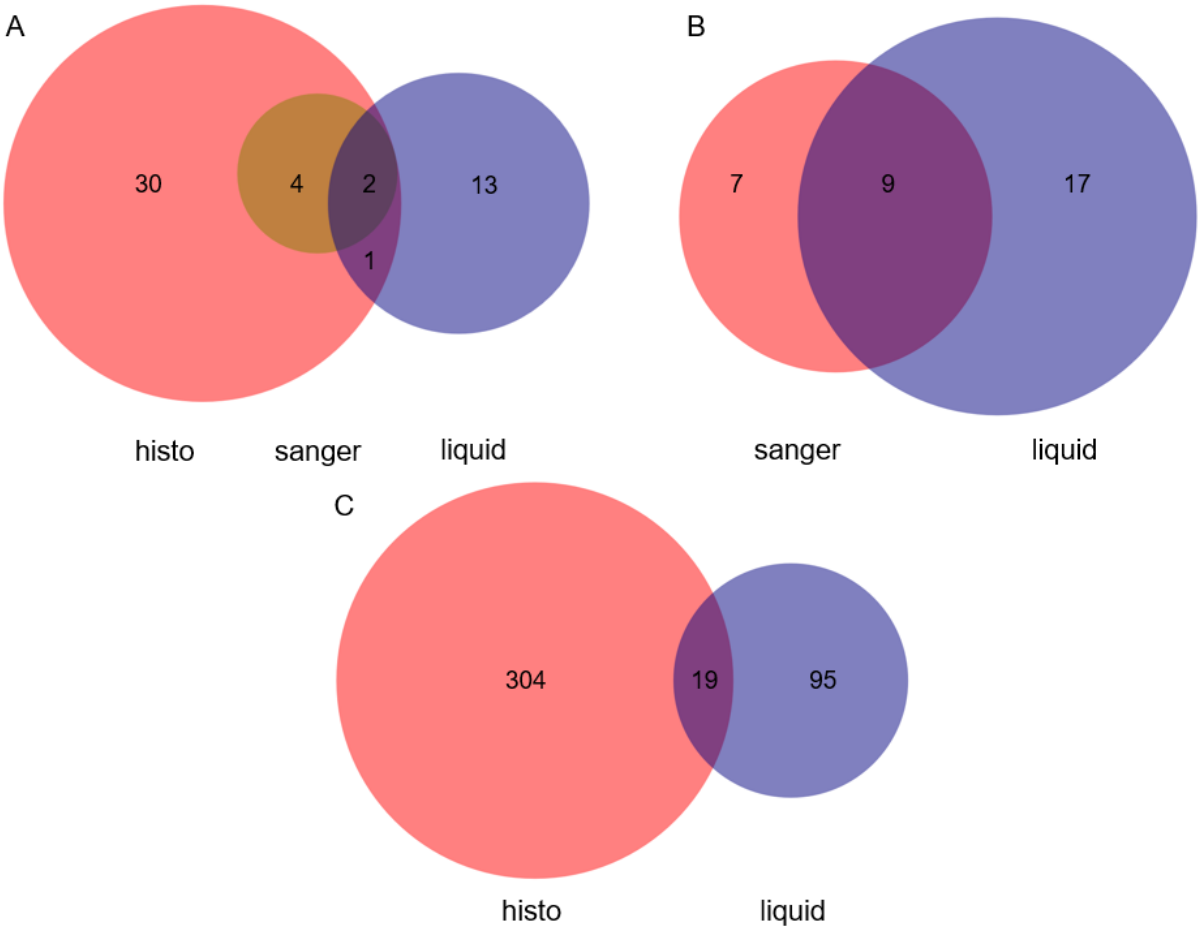

Figure 3-12. Venn diagrams of non-synonymous coding mutations detected with the different methods in tissue and in cirDNA. “histo” refers to the amplicon sequencing performed on FFPE samples. “Sanger” sequencing was carried out routinely on tumor tissue. “liquid” refers to mutation analysis of cirDNA.
(A) KRAS, NRAS and BRAF status obtained by all three methods in 18 patients. 2 KRAS mutations in Exon 4 detected by Sanger sequencing are not depicted as they were not covered by the amplicon panel.
(B) KRAS, NRAS and BRAF status obtained by Sanger sequencing and cirDNA-analysis in 33 patients. Again, KRAS mutations in Exon 4 were excluded.
(C) Number of non-synonymous coding mutations detected by the panel in 19 patients in FFPE and cirDNA samples.

Low concordance of amplicon sequencing between liquid biopsy and FFPE samples

In 10/19 patients at least one mutation found in tumor tissue could be redetected in the blood. The overall concordance was only 4.5% (Figure 3-12C). Mutations that were redetected had higher allele frequencies in the tumor tissue than those that were not found in consecutive liquid biopsies (median 27.40 vs. 1.00; $p<0.0001$). In line with this, their MAF also were higher in cirDNA samples compared to mutations only found in liquid biopsies (median 3.15 vs. 1.07; $p=0.0002$). On average, 14.7 mutations were

found per FFPE sample, whereas in cirDNA samples a mean of 2.41 mutations could be detected.

3.2.4 Mutation analysis of cirDNA in the mCRC-cohort

Mutational patterns in cirDNA are patient-specific and highly recurrent

The mutational profiles in cirDNA of 34 patients with advanced CRC (mCRC-cohort) were examined. Serial plasma samples were available for all patients. Patients received a palliative chemotherapy consisting of a 5-FU backbone with or without oxaliplatin or irinotecan. In addition, most patients received a targeted therapeutic, either targeting VEGF/VEGFR (bevacizumab, aflibercept) or EGFR (cetuximab, panitumumab). Detailed patient characteristics can be found in Table 2-2. Initially, all patients within the cohort showed a response to their therapy, defined as either stable or regressive disease in radiological imaging. 14 patients remained stable or showed partial remission while 20 patients developed disease progression between two consecutive blood samples taken.

In total, 103 samples were sequenced. In 80 samples mutant DNA could be detected. ctDNA-levels ranged from 0ng to 74.89ng/ml plasma (median 0.094ng/ml plasma). Individual mutation analyses are listed in Table 3-1 and illustrated in Figure 3-13. Samples with MAF >2% had significantly higher cirDNA-levels (5.35ng/ml vs 3.82ng/ml plasma; p=0.0371). In general, MAF and cirDNA-level correlated positively, with a PCC of 0.62. Amplicon sequencing of cirDNA revealed on average 2.41 alterations per sample. These alterations were either recurrent (present in serial samples) or sporadic. Alterations occurred at an allele frequency between 0.5% and 96.7% (median 2.25%). Most mutations were found in amplicons of TP53. Overall, a trend towards higher number of alterations in patients with disease progression was observed (Figure 3-13B). The MAF was significantly higher in patients with progressive disease compared to those with stable or regressive disease (Figure 3-13C).

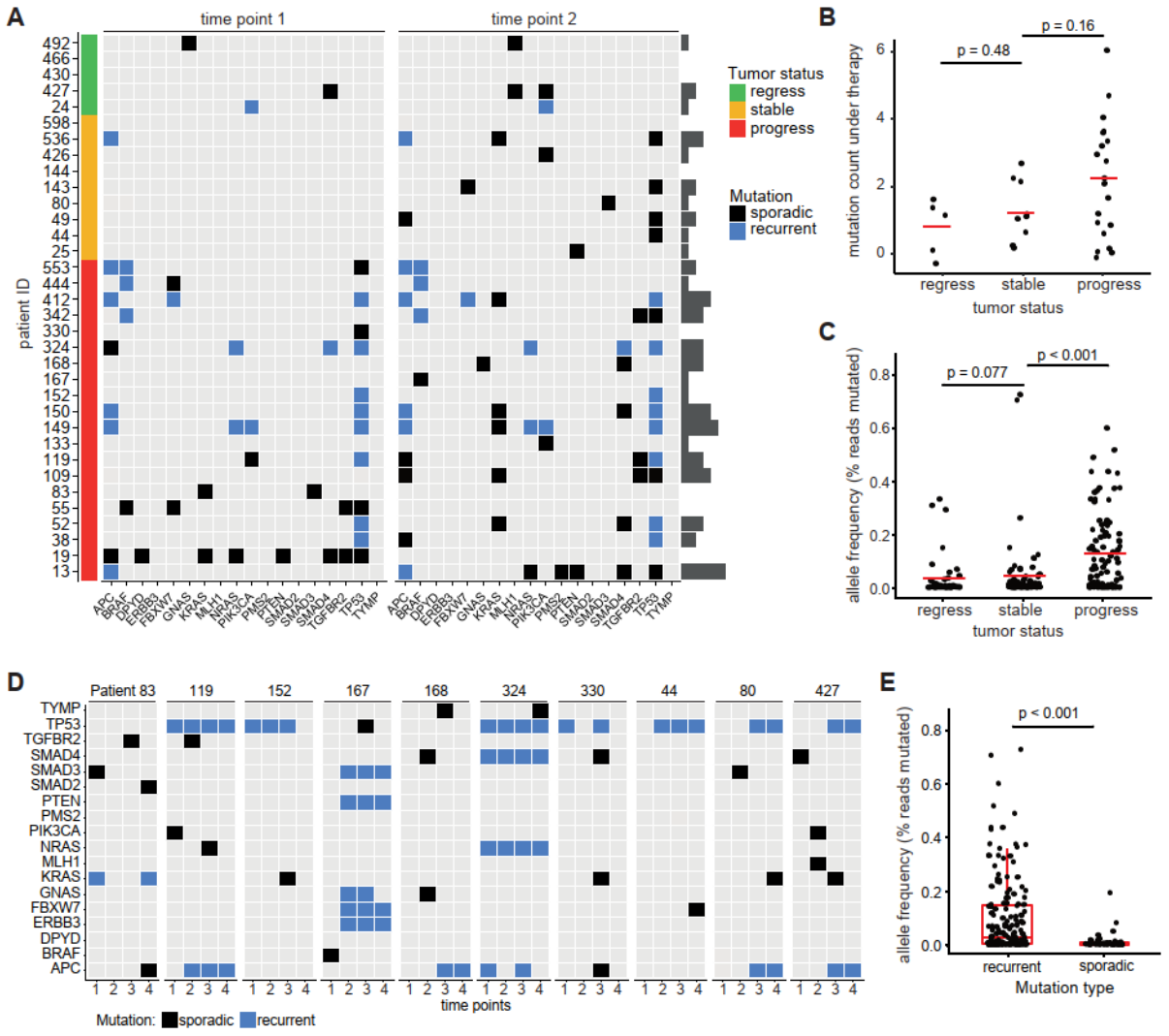

Figure 3-13. Mutational patterns in serial liquid biopsies.
(A) Heatmap of all mutations detected in cirDNA at two subsequent time points while receiving treatment. Disease staging was categorized as regressing, stable or progressive according to radiologic imaging between the two time points assessed. Several mutations were recurrent in both samples, while others were only noted in either time point 1 or 2 (sporadic).
(B) Mutation count under therapy. Tumor status was categorized as regressive, stable or progressive by radiologic imaging at the time the cirDNA sample was taken.
(C) Allele frequency of detected mutations according to tumor status.
(D) Development of mutational patterns over the course of treatment. Four cirDNA samples were taken from each patient at different time points over the course of treatment and sample were analyzed. Several mutations were reoccurring between different time points, while others appeared only at one time point.
(E) Comparison of the allele frequencies of recurrent and sporadic mutations.

Using APC as an example case it could be shown that the distribution of mutations in cirDNA was very similar to the mutational spectrum found in primary CRC tissue (Figure 3-14).[35] In contrast, the mutational congruence of matched tumor tissue and cirDNA samples in this cohort was low (Figure 3-12C).

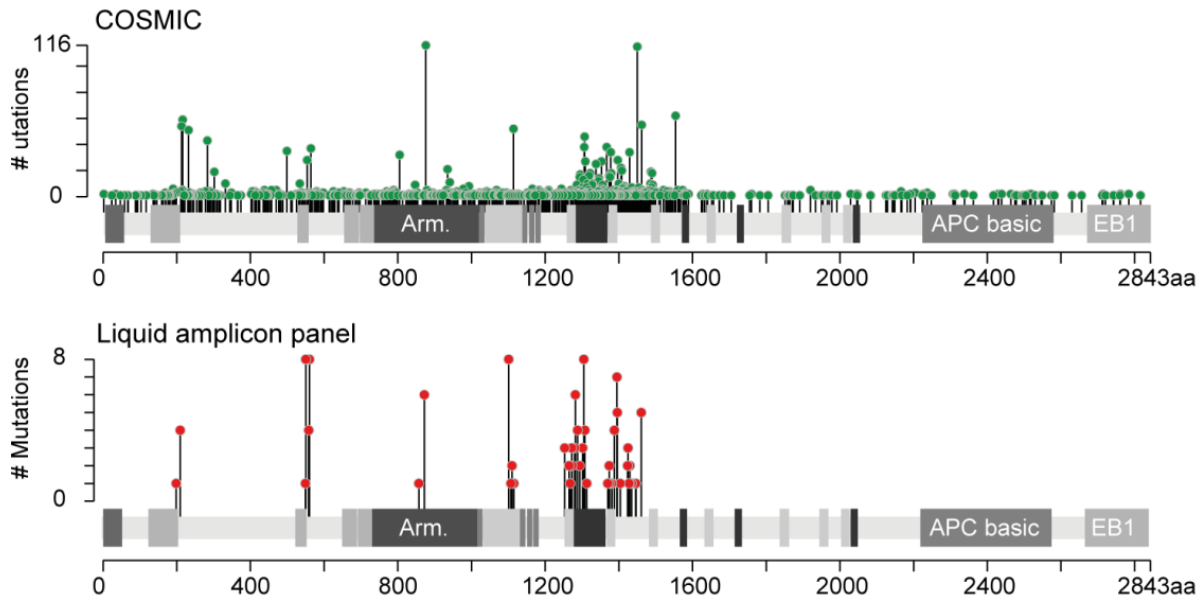

Figure 3-14. Frequent APC hotspots are covered by the amplicon assay. Mutations in APC in the COSMIC database (upper panel) and mutations in APC detected by the amplicon assay in liquid biopsies (lower panel).

For a small group of patients up to four consecutive blood samples over the course of therapy, corresponding to a treatment period of approximately one year, were analyzed. For all patients in this cohort both recurrent and sporadic mutations could be detected (Figure 3-13D). Recurrent mutations were more frequent than sporadic mutations and showed a higher allele frequency (Figure 3-13E). The pattern of recurrent mutations appeared to be distinct for each patient and was conserved throughout months of therapy.

Table 3-1. Individual mutation analysis and liquid biopsy statistics.

Patient ID	Tumor entity	Therapy regime	Sanger	Histo	Liquid	Blood samples	No ctDNA	MAF (%)	ctDNA-level (ng/ml plasma)	Recurrent mutations
13	rectum	FOLFOX + bevacizumab	NRAS Q61L	ND	NRAS Q61L	3	0	7.3-13.8	0.29-0.73	3
19	rectum	FOLFIRI + panitumumab	WT	WT	WT	2	1	0-1.9	0-0.37	0
24	colon	FOLFIRI + cetuximab	WT	WT	WT	3	0	15.3-37.7	0.86-3.23	1
24	lung		KRAS G12C	KRAS G12C KRAS G13A	WT					
25	rectum	FOLFIRI + bevacizumab	ND	KRAS G12D NRAS Q61R	WT	3	2	0-0.7	0-0.028	0
38	rectum	FOLFIRI + aflibercept	WT	ND	WT	3	0	4.0-36.1	0.094-1.87	1
44	colon	FOLFIRI	WT	KRAS G12V NRAS Q61R	WT	3	0	0.7-7.4	0.025-0.62	0
49	colon	FOLFIRI + cetuximab	WT	ND	WT	3	1	0-12.2	0-0.27	1
52	colon	5-FU + bevacizumab	KRAS G12D	ND	KRAS G12D	2	0	2.3-25.5	0.49-13.38	1
55	rectum	5-FU + bevacizumab	KRAS G12V	KRAS G12V	WT	4	3	0-1.0	0-0.068	0
80	colon	FOLFIRI + bevacizumab	KRAS G12D	KRAS G12D BRAF V600E	KRAS G12D	4	0	0.7-6.6	0.028-0.41	1
83	rectum	FOLFIRI + cetuximab	WT	NRAS Q61K	KRAS G13D	4	1	0-1.6	0-0.059	1
109	rectum	FOLFIRI + cetuximab	WT	WT	WT	3	0	1.3-4.5	0.028-0.44	1
119	rectum	5-FU + cetuximab	WT	WT	NRAS Q61K	4	0	0.8-36.6	0.012-1.33	2
133	rectum	FOLFIRI + bevacizumab	KRAS A146T	ND	WT/ND	2	1	0-2.7	0-0.26	0
143	rectum	5-FU + cetuximab	WT	WT	WT	3	1	0-3.4	0-0.076	0
144	colon	FOLFOX + bevacizumab	NRAS Q61R	NRAS Q61R	WT	3	2	0-0.7	0-0.021	0
149	colon	FOLFOX + bevacizumab	WT	WT	NRAS Q61K	2	0	33.5-60.2	1.81-8.11	3

Patient ID	Tumor entity	Therapy regime	Sanger	Histo	Liquid	Blood samples	No ctDNA	MAF (%)	ctDNA-level (ng/ml plasma)	Recurrent mutations
150	rectum	FOLFIRI	KRAS G12D	ND	KRAS G12D	3	0	2.7-43.8	0.074-2.38	4
152	rectum	FOLFOX + bevacizumab	KRAS G12R	KRAS G12R KRAS G12V	KRAS G12R	4	1	0-3.2	0-0.24	0
167	colon	5-FU + bevacizumab	KRAS A146T	ND	WT/ND	3	0	1.1-9.0	0.14-0.27	5
168	rectum	5-FU + cetuximab	WT	WT	WT	4	1	0-3	0-2.11	1
324	rectum	FOLFIRI + cetuximab	WT	WT	NRAS Q61K	4	0	15.3-96.8	1.19-74.89	3
330	colon	FOLFIRINOX + bevacizumab	KRAS G12S	ND	WT	4	2	0-4.9	0-0.36	0
342	colon	FOLFIRINOX	BRAF V600E	ND	BRAF V600E	2	0	37.7-72.7	3.11-50.97	1
412	rectum	FOLFIRI	KRAS G12R	ND	KRAS G12R	3	0	15.8-24.0	0.38-0.97	4
426	colon	FOLFIRI + bevacizumab	KRAS G13D	ND	WT	3	1	0-9.4	0-0.19	1
427	colon	FOLFIRI + cetuximab	WT	KRAS G13D	KRAS G13D	4	0	0.6-49.1	0.021-11.61	2
430	colon	FOLFIRI + bevacizumab	KRAS G12C	KRAS G12C	WT	4	3	0-1.2	0-0.57	0
444	colon	FOLFIRI	BRAF V600E	ND	BRAF V600E	2	0	7.6-11.5	0.44-0.68	1
466	rectum	FOLFIRI + cetuximab	WT	ND	WT	3	2	0-1.7	0-0.36	0
492	rectum	FOLFIRI + bevacizumab	KRAS G12S	ND	WT	3	0	0.6-1.4	0.008-0.044	1
536	rectum	FOLFIRI + bevacizumab	KRAS G12A	ND	KRAS G12A	2	0	1.9-8.6	0.033-0.41	1
553	rectum	FOLFIRI + cetuximab	WT	ND	BRAF V600E	2	0	6.1-33.5	0.16-1.30	2
598	colon	FOLFIRI + cetuximab	WT	ND	WT	2	1	0-1.4	0-0.030	0

“Sanger” sequencing was carried out routinely on tumor tissue. “histo” refers to the amplicon sequencing performed on FFPE samples. “liquid” refers to mutation analysis of cirDNA. WT=wild type; ND=not done. Number of samples and number of those without ctDNA are listed. MAF and ctDNA-level depict the respective range across serial liquid biopsies. Recurrent mutation is defined as mutation detected in two or more cirDNA-samples.

Liquid biopsy allows monitoring the clinical course of CRC patients and reveals resistance mechanisms

Next, the clinical correlation of mutation profiles to the course of treatment in CRC patients was examined. For that, the MAF and ctDNA-levels were correlated with clinical response (Figure 3-15). All four patients of the presented case series initially received 5-FU, folinic acid, irinotecan (FOLFIRI) and cetuximab and showed treatment response (either stable or regressive disease) by the time of the first liquid biopsy. P119 had achieved remission and her therapy was deescalated to 5-FU + cetuximab. Two novel mutations occurred in the cirDNA upon disease progression. The respective MAF closely matched the radiologic extent of disease and the CEA-values (Figure 3-15A). Analogously, in P324 (Figure 3-15B), P24 (Figure 3-15C) and P553 (Figure 3-15D) MAF increased upon tumor progression, while therapy response was associated with stable or falling MAF. CEA levels did not correlate with re-staging results and MAF in P553.

In all these cases an increase in MAF was closely associated with radiological disease progression. Overall, considering only the 22 patients with recurrent mutations, changes in MAF correlated with clinical response in 25/33 cases. Stable disease was associated in 5/7 cases with an increase in MAF and eventually all patients developed a progress. Disease progression coincided with a rise in MAF in 20/24 cases, i.e. a sensitivity of 83% for discovering disease progression. For 12 patients in this cohort ctDNA analysis did not yield enough results to make plausible comparisons to therapy response. These results highlight the potential of cirDNA as an individual biomarker.

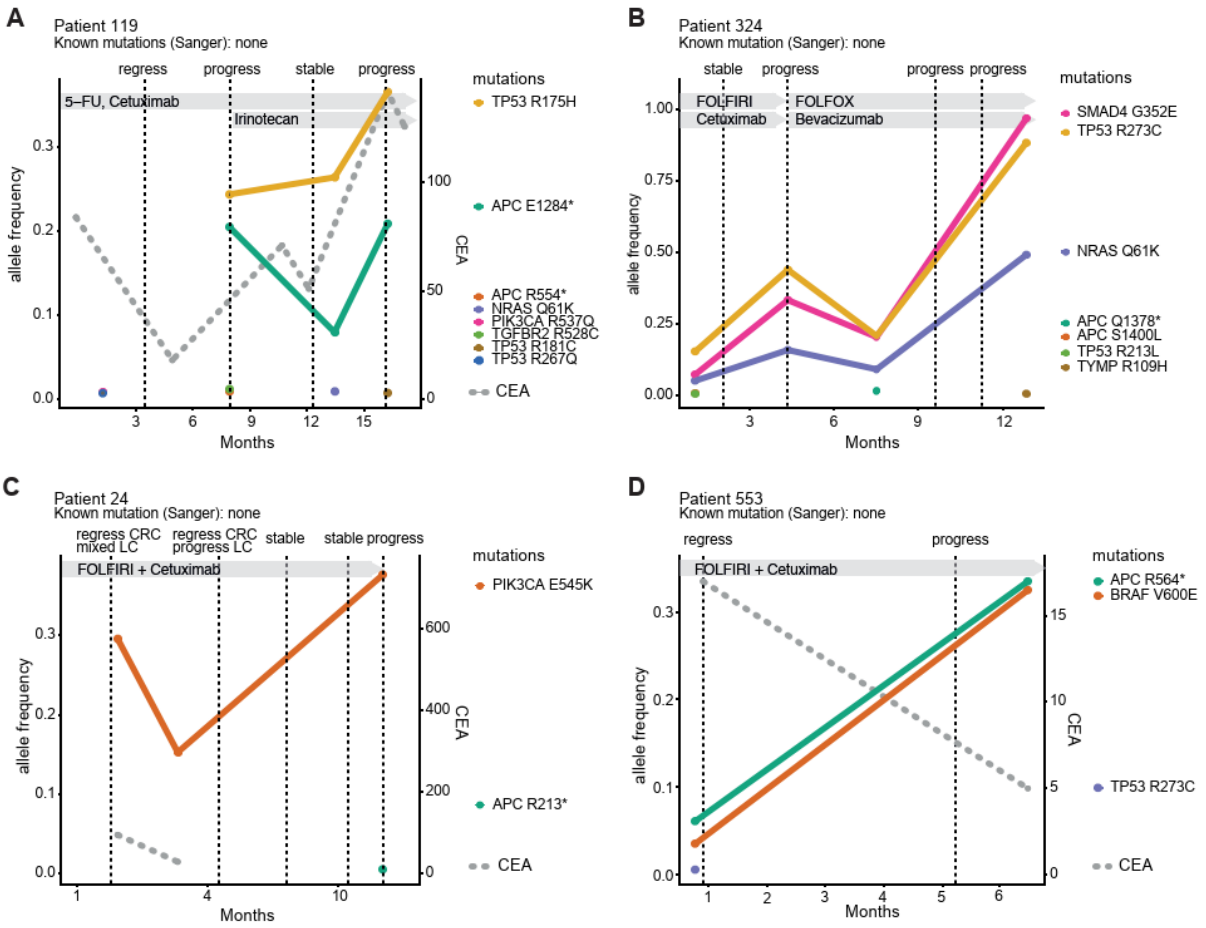

Figure 3-15. Monitoring the clinical course of CRC patients by amplicon sequencing of cirDNA. **(A)** P119 had sustained response of her metastasized rectum carcinoma to the maintenance regimen (5-FU and cetuximab) for almost two years and very low concentrations of cirDNA. When disease progression was first noticed, two mutations with >20% MAF arose, while cirDNA-levels were <1ng/ml plasma. Re-escalation to FOLFIRI + cetuximab led to stable disease in the next re-staging and progressive disease another three months later. Accordingly, the MAF remained constant/decreased and eventually doubled, while the cirDNA-level rose from 2.49ng/ml to 3.64ng/ml plasma. Findings were in line with changes of CEA.
(B) P324 had a local recurrence of rectal cancer and hepatic filiation. After initial response, the patient was diagnosed with lung metastases and a corresponding rise in MAF of three recurrent mutations could be noticed, thereunder an NRAS mutation that had neither been detected by Sanger sequencing nor amplicon sequencing of the primary tumor. Her therapy regimen was changed to FOLFOX + bevacizumab. In the next blood sampling, the MAF declined, however two month later in the next re-staging examination the patient showed a metastasis in the pancreas and eventually she had an extreme increase in both cirDNA-level and MAF (calculated ctDNA-level 74.9ng/ml plasma) and died within two weeks. CEA levels were not available.
(C) P24 was diagnosed simultaneously with non-small cell lung cancer and hepatic metastasized colon cancer. Each of the tumor manifestations were histologically proven. The colorectal cancer was RAS wild-type, the lung cancer was RAS-mutated. Nevertheless, the patient received FOLFIRI and cetuximab. While treated, a PIK3CA mutation that was also identified by amplicon sequencing in material from the colon cancer, sank in allele frequency before radiologic imaging proved regression of the colorectal cancer (and progression of the

lung cancer). Upon progress of the colon cancer the MAF jumped up. ctDNA-levels changed from 3.24ng/ml to 0.86ng/ml and 2.36ng/ml plasma and like the MAF reflect the status of colon cancer well.
(D) P553 was treated for metachronous liver metastases of RAS/RAF wild type rectal cancer with FOLFIRI and cetuximab. After initial tumor regression, allele frequencies of both the APC and BRAF mutation increased. ctDNA-levels increased eightfold upon tumor progression from 0.16ng/ml to 1.3ng/ml plasma, whereas CEA-levels sank. The BRAF mutations was not detected by Sanger sequencing in primary tumor tissue seven month earlier.

In six patients formerly RAS and RAF WT treated with cetuximab new mutation known to confer resistance to anti-EGFR agents arose (Table 3-1). In P83, P149, P324 and P553 those mutations could be detected in serial samples. P149 already had progressed under cetuximab and was changed to FOLFOX + bevacizumab before first sampling. In P427 the same mutation was detected by amplicon sequencing of FFPE tissue. P119 displayed a rare NRAS mutation at one time point. All patients had progressive disease in the next four months after mutation detection. The progression-free survival (PFS) was shorter compared to the other patients of the cohort (median 3.0 vs 5.3 months; p=0.0055).
Other mutations in genes associated with MAPK signaling and currently not implemented in primary diagnostics include PIK3CA, PTEN and ERBB3. P24, P149 and P426 displayed recurrent PIK3CA mutations. P167 showed a common truncating PTEN mutation and an ERBB3 mutation. P426 and P167 had a PFS of 3.5 and 3.8 months, whereas the colon cancer of P24 first showed a partial response with a progress after 12.5 months.
The attained knowledge about RAS and RAF mutation status could have facilitated the decision to discontinue treatment earlier in at least five patients on the basis of accepted therapy guidelines,[68] giving them a chance to respond to a different regimen and sparing them from unnecessary side-effects.

ctDNA shows superior properties as blood biomarker compared to cirDNA
Interestingly the correlation between ctDNA-levels and blood parameters in mCRC patients was stronger compared to cirDNA with the same blood parameters. Strong correlations were observed between ctDNA-levels and LDH, CRP, CEA, CA19-9 and albumin (Figure 3-16). Regarding cirDNA-levels of these samples, a correlation existed only with LDH, CRP and albumin. The unique association between serum tumor biomarkers and ctDNA-levels highlights their specificity as potential tumor markers.
Along with assertions about individual mutation profiles and their implications in treatment response ctDNA acts as a biomarker with multiple applications.

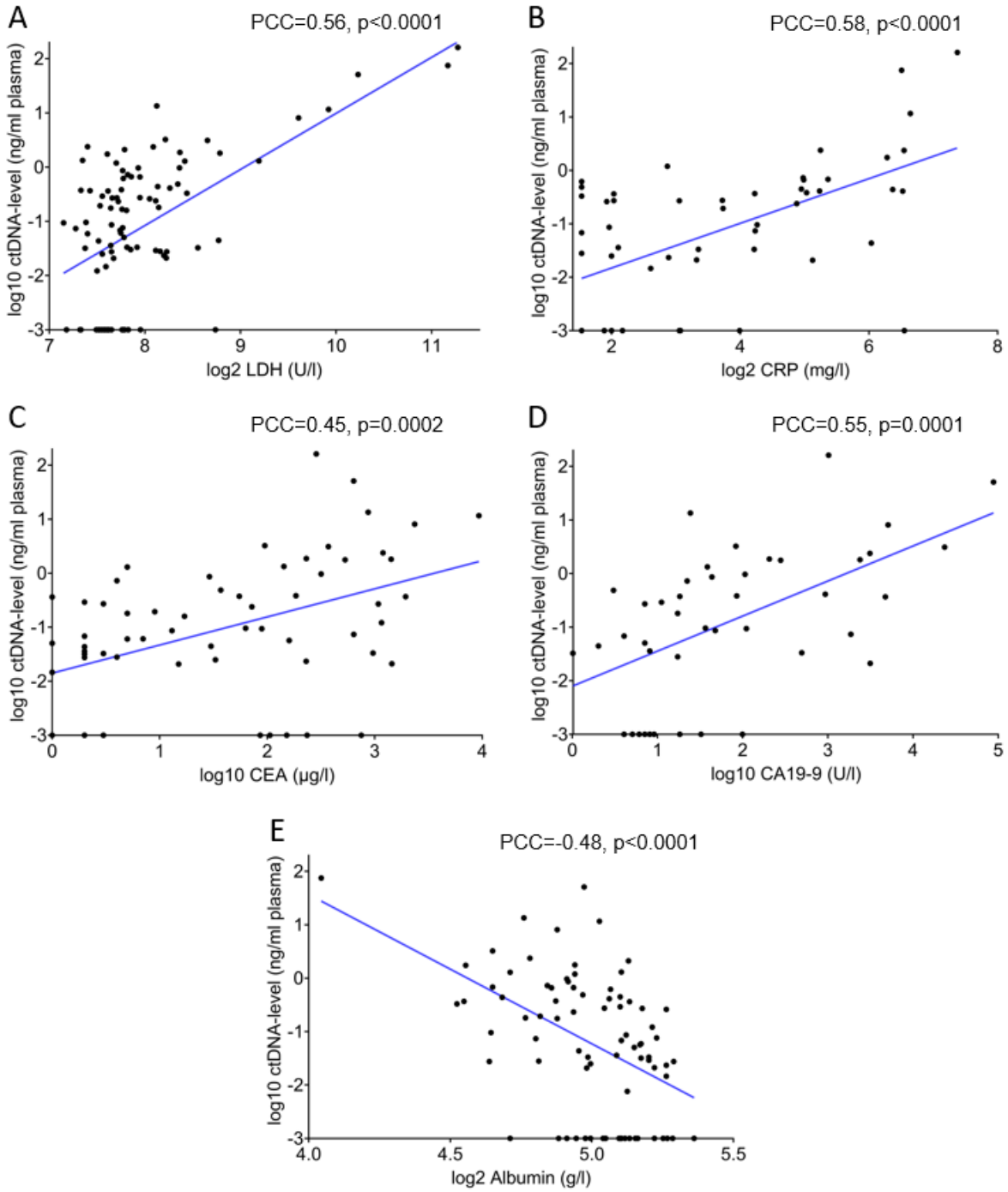

Figure 3-16. Correlations of blood parameters with ctDNA in patients with advanced CRC.
(A) Association of ctDNA-levels with LDH-levels. The correlation was not limited to high LDH values.
(B) Association of ctDNA-levels and CRP. The assay for CRP measurement had a lower detection limit of 2.9mg/l. This limit was taken as default value for samples with no detectable CRP.
(C) CEA vs. ctDNA-levels. Rising CEA-levels were accompanied by rising cirDNA-levels.
(D) Association of CA19-9 and ctDNA-levels.
(E) Association of albumin and ctDNA-level. A negative correlation was observed.

4 DISCUSSION

4.1 Multiplex amplicon sequencing as detection method for ctDNA

The method described in this thesis is a highly flexible experimental and bioinformatics pipeline to analyze custom defined hotspots in cirDNA by multiplexed amplicon sequencing. This method for cirDNA analysis was shown to be feasible in clinical settings.

Until now, various approaches have been developed to characterize cirDNA either looking at specific mutations or examining multiple hotspots and genes.[77] A major limitation to their broad clinical application is the genetic landscape of cancers. Especially mutations that only occur in selected cancer types can represent actionable mutations, albeit they are not covered by commercial gene panels. Examples include mutations in RNF43[155] or PMS2,[156] which could indicate subclones that make the tumor responsive to specific therapies such as immune therapy. Hence, a key challenge is to customize the analysis of cirDNA for specific tumor entities. On that front, this study supplies a methodological scaffold that enables the design and consecutive analysis of custom amplicon panels using open source bioinformatic tools. Applying this method, an amplicon panel tailored for colorectal cancer was designed, including hotspot mutations in oncogenes and tumor suppressor genes, as well as rare mutations in genes relevant for DNA repair and drug metabolism.

4.1.1 Integrating deep sequencing into routine management

Targeted amplicon sequencing offers an efficient approach to detect a considerable amount of common mutations. In comparison to WGS and WES a very high depth (>1000 reads) per hotspot can easily be reached, thus enabling detection of low-frequency variants in heterogeneous tumor samples and in cirDNA samples, with low overall amounts of ctDNA.[157,158]

In addition, the low overall number of base pairs sequenced makes it possible to analyze several samples in parallel on platforms such as the Illumina MiSeq or HiSeq. The comparably low costs have the potential to allow future routine application.

4.1.2 Low quality sequencing material - cirDNA and FFPE tissue

FFPE-tissue and cirDNA pose analytical problems. Formalin fixation and paraffin embedding is the widely used standard for biopsy conservation, because of the simple and low-cost storage possibility.[159]

cirDNA and isolated DNA from FFPE samples are highly degraded into short fragments, mostly smaller than 200bp. Formalin fixation also causes covalent binding of nucleic acids with proteins.[159] Sequencing technologies need to be adapted to these conditions, still less robust results and difficulties in data interpretation arise.

4.1.3 Advantages and limitations of amplicon sequencing

Formation of primer dimers is a common PCR artefact, which limits PCR usability by decreasing the amplification of the target fragments and heisting sequencing depth. These problems often increase when multiplex PCR is applied. One of the strategies to diminish the negative effects is to make use of the length differences. Primer dimers have a length of about two times the length of a single primer. In this experimental approach dimers were about 90bp long − as a primer in the first PCR had a length of 44-52bp. Target products after the first PCR were at least 150bp long. Thus, via purification with paramagnetic metal beads efficient size selection could be performed. To a smaller extent this could also be achieved with the second clean-up.

Another way to enrich target products is the segregation of DNA fragments on agarose gels. After electrophoresis the desired size area can be cut out. This method needs multiple steps, more DNA input, more hands-on time and is only feasible for a small number of samples.[160] In contrast, DNA purification with AmpureXP beads could be semi-automated, significantly decreasing the hands-on time per sample. Other PCR purification protocols do not allow to adjust the cut-off limit and are therefore only suitable to wash out fragments <100bp, i.e. primers, salts and to a certain extent also primer dimers.[161] Automatization and standardization of the workflow in a 96 well format not only led to a drastic labor reduction but also minimized potential error sources.

The other strategy is to prohibit the formation of primer dimers in the first place. In order to achieve this, special attention is needed when designing a multiplex PCR reaction. By the use of open source software, like primer3 and MultiPLX 2.1,[143,144] primer combinations can be checked for high probability of dimer formation. The introduction of a hairpin structure into the primer sequence as described by Stahlberg et al.[146] could

further decrease the amount of primer dimers. They also applied a preamplification protocol as first PCR step. The drastic reduction of primer dimer through the latter modification could be confirmed by the experiments of this thesis.

Multiplex PCR causes inequal amplification of different targets due to varying primer-pair performance. Final sequencing analysis will struggle with manifold differences in amplicon coverage. To call rare variants as true mutations a certain coverage is needed. On these grounds, a tight distribution of amplicon coverages will need fewer sequencing reads to cover all amplicons sufficiently. In the nested PCR protocol applied here the second PCR uses only one primer-pair and thus it can be assumed to amplify all fragments evenly. This means to narrow any variances the first PCR must be optimized. By running only 18 cycles and assuring proper annealing of primers and their respective target, amplicon variance could be lowered down to 41-fold. A method termed as SiMSen-Seq recently published by Stahlberg et al.[146] runs just three preamplification cycles and reported read depths that varied only 8-fold. In this protocol 36 PCR cycles in total were applied, which is slightly more than in other established protocols (27-33 cycles).[146,162] Eventually libraries with a low fraction of nonspecific products were obtained and ensured high quality sequencing.
Performance of multiplex assays cannot be predicted with high confidence using currently available software and a few primer pairs had to be discarded because they yielded no functional amplicons. Experimental validation of new multiplex reactions will be necessary and is doubtless the most time-consuming developmental step.

In this assay only a few hotspots of the MMR genes were targeted, and epigenetic silencing was not examined. The included MSH6 hotspot proved to be problematic to examine within this assay due to a long cytosine repeat. Thus, results had to be discarded in several cases. As most MMR genes lack definite hotspots, the utility of amplicon sequencing is limited. Another solution might be to gather the TMB as most MSI-tumors are hypermutated and the TMB is of relevance for immunotherapeutics.[28,163]

As one haploid genome is weighing ~3.3pg roughly 151 copies of each gene could be present in 500pg of DNA.[164] In the case of samples with very low MAF and/or mutations of tumor subclones the sensitivity of the assay could be limited, due to the absence of a mutated copy in the respective reaction. This might have caused missing of low-frequent alterations in some samples. Other reasons for the lack of ctDNA could be

amplification errors, poor DNA quality, false negatives in variant calling or that mutations were not covered by the panel. Some tumor suppressor genes like PTEN and TP53 can be homo-deleted in CRC and such large deletions are not recognizable by this assay.[33] Chromosomal aberrations, gene amplification and epigenetic modifications are also not detectable with an amplicon-based approach. Last, in some samples no tumor DNA could be present after all.

The mutational profile of a tumor is unknown and its anticipation is especially difficult for tumor entities with low mutational burden or heterogeneous and low frequent hotspots, such as thyroid cancer and most sarcomas.[41] Customizing individual panels based on sequencing results from tissue biopsies or surgical specimens is a highly specific solution but also very labor-intensive. In the case of tracking minimal residual disease and recurrence it delivered promising results.[77] The presented method can be used to tailor panels to those specific questions.

All methods for cirDNA analysis rely on PCR-techniques for amplification and consecutive sequencing, the consequence being that errors can always be introduced. It is crucial to keep these false mutations very rare and to identify them in order to increase the positive predictive value of the respective assay.

An effective method to reduce these technical errors is the use of unique molecular identifiers (UMI). Barcoding with UMIs enables tagging of single fragments and permits background noise reduction and higher sensitivity. UMIs are short random nucleotide repeats integrated into the primer tail.[162] For the analysis of cirDNA, methods have been developed to introduce UMIs during the first PCR amplification. These approaches, for instance Safe-SeqS[162] or SiMSen-Seq,[146] have been shown to suppress PCR errors and enable precise counting of mutant allele copies. A drawback of using UMIs is that a vastly higher sequencing depth is required to discriminate between different UMI families, which will increase the costs of the method. Optimal implementation needs at least 240000 reads per target.[146] However, as sequencing costs will further decrease, the integration of UMIs in our approach will be feasible.

With years of experience in next generation sequencing (NGS) a gold standard for data analysis is still non-existent. So far, multiple bioinformatic tools for NGS data have been developed leading to varying results.[165] cirDNA as new input material highlights the urge of specialized pipelines for sample processing and data analysis in the context

of amplicon-based methods. Robust bioinformatic pipelines as applied here are vital for sequencing analysis and help discarding false positives.[166] Genomic DNA is as a serious confounder, causing to overlook true mutations and detect false positives in normal WBCs. Additional sequencing of WBCs can identify those mutations which are due to clonal haematopoiesis.[167]

Amplicon sequencing is a trade-off between sensitivity to find rare mutations on the one side and test a considerable number of hotspots on the other side. ddPCR or other single-locus techniques can detect single mutant fragments in plasma, reported limits reach 0.0001%.[94] Detrimental, however, is the fact that for each nucleotide alteration a specific assay must be designed, confining their application to very prominent hotspots. On the other side, current sequencing costs only allow low-depth WES and WGS, which makes it impossible to detect mutations <5% allele frequency.[77] In this intermediate zone, targeted sequencing as established in this thesis, is the optimal tool and in spite of several limitations, the amplicon-based assay proved reliable in detecting cancer mutations in cirDNA as shown by multiple quality control experiments.

4.1.4 Perspectives

Improvements of the assay should focus on higher multiplexing, as thereby the overall input amount can be reduced or the amount per well can be increased. Up to 30-plex has been shown to be possible analyzing cirDNA.[146] By adding other suitable targets more mutations will be identified and the fraction of tumors with at least one alteration covered will rise. Possible hotspots in CRC are for example KRAS A146, CTNNB1 S45, AXIN2 G665, TCF7L2 K485+R488, RNF43 G659, ACVR2A K437. Also, covering the whole TP53 and APC genes with the highest mutation rates and scattered mutation profiles can further enhance the coverage.[33,168]

Upcoming data state that ctDNA fragments are shorter than regular cirDNA fragments,[86,87] which is why development of shorter amplicons (~100bp) is preferable to ensure proper amplification and prevent loss of mutant copies. Together with *in silico* methods sensitivity of mutation detection can be further increased.[169]

Despite all elucidated improvements, mutant DNA will not be detected by amplicon sequencing in some samples. Here, extremely sensitive approaches which look at single-locus alterations can close the gap.[94] On the other side, ways to predict suitability of patients for liquid biopsies should be developed. Extrapolating the cirDNA and ctDNA-level as shown here with the help of laboratory markers may help to identify those samples. Similarly, a sensitivity prediction score developed by Jovelet et al.[170]

used clinical information and blood parameters to select patients for whom cirDNA could function as surrogate marker.

Biological consequences from mutations may vary, classifying genes as WT and mutant is not enough to account for the impact on tumor evolution. Evermore specific mutations are associated with specific phenotypes. For example, different PIK3CA mutations were described to have different oncogenic impact or influence on anti-EGFR treatment,[18] and non-V600E BRAF mutations can impair the kinase function of BRAF and have been associated with favorable prognosis.[55] Future investigations should focus on the differences between specific alterations, as well as in the context of mutational profiles. By monitoring the mutational profile, CRC-tumors can be sorted for example into consensus molecular subtypes as proposed by Guinney et al.[171] Stratification of this or similar classification systems can build the basis for individualized therapy interventions.

Immunotherapy is a new kind of cancer therapy and is now tested in CRC patients. In principle the body's own immune system is empowered to fight "foreign" cells, while so far almost all treatment strategies have tried to target the tumor itself. Mutated DNA-fragments are presented on the tumor cells by the major histocompatibility complex I (MHC-I) and are recognized by T-cells as neoantigens, causing an immune response.[172] It is hypothesized that due to a high TMB more neoantigens are presented on the cells, making it easier for the immune system to react.[163] Supporting this theory is the fact that the hypermutated MSI-CRC contain prominent lymphocytic infiltrates.[173] Not only could this explain the survival benefit of MSI-tumors, but also why they show such high response rates to immunotherapy.[163] Current immunotherapeutic agents target and inhibit either the cytotoxic T-lymphocyte-associated antigen 4 (CTLA-4), programmed cell death protein 1 (PD-1) or ligand (PD-L1) and promote T-cell activation. These proteins normally function as immune checkpoints that maintain self-tolerance and protect tissues by inhibiting the antigen-specific T-cell response.[172] Checkpoint-inhibitors were first successfully used in malignant melanoma and lung carcinoma. These two entities bear very high TMB and their clinical response to immunotherapy correlates with the TMB.[28,41] This concludes why it is relevant to determine the MSI-status, as it may also have therapeutic consequences.

Until now, knowledge about ctDNA origin and its biology just scratches the surface. Profound knowledge about fundamentals of cirDNA dynamics like release, clearance

and respective differences of ctDNA dynamics would help implementation of liquid biopsies into clinical routine and need to be addressed in future experiments.
Other circulating cell-free nucleic acids, CTCs or macromolecules also carry information about the tumor and are noteworthy. Other methods such as WGS, WES, methylome and transcriptome analysis can yield complementary data. Therefore, tests focusing on those changes and sources offer huge additional benefits.[174] Regarding the limited amount of tumorous material in the blood stream intelligent solutions are needed to maximize the information gain and spare unnecessary loss of ctDNA.
This variety of information that can be collected and the consequences on disease progression need to be systematically investigated. The benefit of liquid biopsy-guided treatment decisions needs to be proven in randomized clinical trials.

4.1.5 Conclusions

All in all, this amplicon assay is a very simple and cost-effective but also robust mean for ctDNA amplification and detection, and the overall low amount of DNA needed stands out as it enables analysis of most cirDNA samples. The freedom to adjust the assay to varying demands is another eminent advantage. The results of the mCRC-cohort suggest that amplicon sequencing of cirDNA could be practicable in clinical routine for patients with late-stage colorectal cancer and allow real-time monitoring of the disease. Following studies should aim to reproduce and validate these results in larger cohorts and find out whether they can be transferred to other tumor entities and earlier stages of disease.
Technical improvements like the introduction of UMIs to increase sensitivity and specificity of amplicon sequencing need to be further implemented and methods allowing different analyses of liquid biopsies need to be developed.

4.2 cirDNA and ctDNA and their relevance as novel blood biomarkers

Quantification of cirDNA and ctDNA and individual sequencing results allowed multiple derivations about the status of the patient’s disease to be made. The mutational profile of individual tumors can be monitored in serial liquid biopsies and new information about the progress of the disease can be obtained.
Over the past years, enormous progress has been made in the field of liquid biopsy. Recently, two large studies have thoroughly dissected the mutational landscape of cirDNA in CRC and other tumors.[175,176] Both studies demonstrated that the

composition of mutations is highly similar between cirDNA and the corresponding primary tumors. Therefore, as shown by this thesis and other studies, detection and analysis of ctDNA can be used as a biomarker to monitor tumor progression, treatment response and disease recurrence and may direct future therapeutic approaches.[92,126,129]

4.2.1 cirDNA and ctDNA-levels and their significance

As previously observed and confirmed in this cohort, cirDNA-levels correlate with clinical stage, therapy setting, LDH, CEA and albumin levels.[170,177] Especially high LDH-levels where associated with high cirDNA-levels in multiple gastrointestinal tumors and also a correlation with CRP was noticed. Elevated cirDNA-levels during infection have been reported.[178] The overall tumor burden was not collected from imaging results in this study but has also been linked both to CEA-levels and cirDNA-levels.[177,179] Altogether, the correlations were weak and not sufficient to lead to consequences for individual patients. The vast number of factors influencing the amount of DNA circulating at one time point hinders clear validity.

ctDNA could be detected in at least one of the liquid biopsies of every patient and in 78% of all samples. These numbers are comparable to reported frequencies in CRC (80-90%).[92,176] Mutations found by this assay highly match reported CRC profiles.[33,175,176] TP53 had the highest mutation rate, followed by APC deviating from large sequencing studies. This might be due to the lower overall coverage of APC in our panel as it is a large gene. Higher mutation rate for TP53 have been reported for advanced stages of CRC.[180]

When comparing ctDNA-levels or MAF to clinical and blood markers similar correlations as for cirDNA were observed. The relations were, however, considerably stronger. The MAF was higher in progressive disease, consistent with previous findings,[124–126] and higher MAFs have been associated with worse prognosis and shortened PFS.[92,121,122] Here, a rise in MAF frequently coincided with a disease progress or even preceded a radiologic progress. Correspondingly, Misale et al.[97] observed a rise in MAF several months before progress could be confirmed by radiologic imaging. These findings reinforce the utility of ctDNA as prognostic marker.

4.2.2 Conventional tissue biopsy

As primary diagnostic tool, tumor biopsy remains the gold standard at present, also because definite assertions about tumor origin and grading can only be made by histological sampling. However, the invasive acquisition represents a major barrier, that forbids numerous re-biopsies, as they can have serious complications.[181] Tissue biopsies comprise a large spectrum of quality. This depends primarily on the specimen size and the tumor cellularity (percent tumor). Besides, formalin-fixation worsens DNA quality, rendering samples in some cases useless for specific molecular analyses. Taken together, biopsy failure rates range between <10% to more than 30%.[182]

Even if adequate material was obtained, sampling biases arise from tumor heterogeneity. This applies to intratumoral (distinct genetic profiles of separate tumor areas) and intertumoral heterogeneity (differences between primary tumor site and metastases).[183,184] A biopsy will only represent a small tumor section and no information about the rest of the mass or other tumor sites will be gathered. On the one hand, large sequencing studies have confirmed the concordance of genetic alterations in cirDNA to primary tumors.[175,176] On the other hand, ctDNA deriving from different metastatic lesions have been detected and thus ctDNA is likely to be a molecular marker of the overall disease and multifocal clonal evolution.[92,139,140].

More than half of RAS and RAF mutations as assessed by Sanger sequencing could be redetected and in 10 of 19 patient's, tissue and liquid biopsy had a matched mutation. In contrast panel results of all mutations displayed low concordance between tissue and liquid biopsy. So far, reports showed medium to high concordance (39-93%) between matched tissue and liquid biopsies.[170,180,185–188] Kaiseki et al.[188] noted high discordance of initial tissue and later liquid biopsies in melanoma patients due to intermediate treatments and time spans over one year. In the present mCRC-cohort the majority of patients had received multiple therapies before the first blood sampling and often years lay between biopsy and blood sampling. Also, little is known about tumor dynamics under therapy and the fate of tumor subclones. Here, deep sequencing of histological samples revealed multiple low frequency mutations representing tumor subclones, which might have vanished later due to a selection disadvantage under therapy. Even assuming they remained present until cirDNA sampling their low frequency would have prohibited their redetection in most cases in the context of low ctDNA proportion (68% of samples had a MAF <10%) as the detection threshold for variants is 0.5% in this assay. This detection limit is comparable

to other targeted assays.[77] In a few samples failure of the assay might have prevented redetection.

Selective pressure and clonal evolution could also account for mutations only found in liquid biopsies and explain mutational changes in serial blood samples. Together with similar observations these data point out the benefit of cirDNA analysis to investigate tumor heterogeneity.[170] In line with previous reports conserved mutational profiles could be found in serial blood samples.[99,125,127,186] Recurrent mutations had higher frequencies and where more often found in the initial tumor biopsy. This stem mutational signature could represent conserved tumor clones. This is especially interesting for application in earlier stages, as it could be used to easily monitor minimal residual disease and recurrence. First promising results, which identified patients at high risk of recurrence by testing for postoperative ctDNA-levels, have been published.[129] McDonald et al.[189] could prove the feasibility of ctDNA analysis in a neoadjuvant setting, identifying patients with incomplete therapy response.

The group of patients with non-overlapping Sanger and liquid biopsy sequencing results had lower ctDNA-levels and fewer recurrent mutations. As the assay could not detect mutations in KRAS and NRAS in multiple plasma samples, clonal evolution might have led to the disappearance of these mutant subclones. These patients might profit from a therapy with anti-EGFR agents. Supporting evidence for this rationale was raised by Khan et al.[190]

4.2.3 Serum tumor biomarkers

Current serum tumor markers, such as CEA and CA19-9 in CRC, lack sensitivity and specificity in diagnostics.[191] Only 70% of colorectal tumors release CEA,[191] which is considerably lower than detection rates of ctDNA and the long half-life impedes real-time monitoring.[192] So far, there have been only small studies comparing ctDNA to CEA, which have found similar correlations as shown by this study,[177,179,180] and also cases in which ctDNA could detect a relapse earlier.[127,186,193]. In patients with breast and ovarian cancer, ctDNA proved to be superior as prognostic predictor compared to standard serum biomarkers.[100,127]

4.2.4 Monitoring therapy response - Individual assertions from mutational profiles in cirDNA

ctDNA can provide information about the individual mutation profile over time, a characteristic that current tumor biomarkers are lacking.
The mutational profile of a given tumor creates a unique biological signature that is highly specific compared to other tumors and normal tissue. Single mutations only found at one time point, however, might be due to clonal hematopoiesis. Mutations in TP53 and KRAS have been found at low levels in blood plasma of healthy individuals.[132,166,167] In a subset of these “healthy” individuals the mutation detection preceded a tumor diagnosis, stating that healthy individuals with mutated fragments in liquid biopsies might be at increased risk.[132] The detection of multiple mutations in tissue and/or serial blood samples will ensure the high specificity of ctDNA as tumor marker.

Emergence of resistance-conferring mutations under anti-EGFR therapy could be observed in this cohort. It was associated with poorer response, analogous to previous findings.[92,190] To increase recognition of resistance, testing for other resistance mechanisms, like gene amplifications, should be implemented. These findings can help physicians to discontinue therapy - sparing patients from side effects - and offer other therapeutic approaches, which may improve prognosis. To what extent patients can benefit from such interventions needs to be studied. Extensive characterization of *de novo* mutations under therapy is needed and will identify more actionable mutations in the context of anti-EGFR treatment but maybe also to classic chemotherapeutic agents. Some genes may contribute to the resistance to 5-FU or platinum compounds if mutated. Candidate genes are for example SMAD4 and FBXW7.[64,194] Therapeutic consequences of their inactivation in CRC need to be addressed to define personalized therapy approaches. The role of 5-FU metabolism genes is shadowy but can theoretically play a central role in therapy response. Mutations in DPYD and TYMP have been found in blood samples of this cohort. Albeit, these single cases do not allow foreseeing their functional relevance. As 5-FU is a standard therapeutic against multiple cancers the interest to learn more about possible resistance mechanisms remains great. A major challenge of future projects will be the functional annotation of these rare mutations and the understanding of their impact on tumor biology and therapy resistance.

4.2.5 Conclusions

This data showed the utility of ctDNA as prognostic and predictive biomarker and its superiority to the measurement of cirDNA alone. Its serial determination offers a variety of information with implications for the patient and his subsequent therapy.

ctDNA represents a blood-based biomarker that can exceed the prognostic and predictive value of current serum tumor markers in advanced stages and maybe also in the context for early diagnosis, recurrence detection and prediction.[129–131]

Future larger studies need to integrate ctDNA analysis in a standardized clinical setting and address its predictive and prognostic role in CRC, in different stages of disease and other tumor entities.

Modern oncologic diagnostics consist of multiple tools and tests that in the end intertwine and draw a picture of the patient and his disease. Clinical examinations, imaging studies, histologic sampling and various blood parameters are eminent foundations of modern diagnostics. Liquid biopsy must be one item in this system, and it will allow higher resolutions on the way to personalized medicine.

5 SUMMARY

cirDNA has long been known to be present in the blood. In cancer patients, levels of cirDNA are higher and they rise further in patients with advanced stages of disease. Importantly cirDNA in cancer patients contains mutated fragments originating from the tumor. As monitoring the mutational patterns of solid tumors during cancer therapy is an unmet need in oncology, analysis of mutations in circulating DNA offers a non-invasive approach to detect mutations that may be prognostic for disease survival or predictive for primary or secondary drug resistance. A main challenge for the application of cirDNA as a diagnostic tool is the diverse mutational landscape of cancer. Here, a flexible end-to-end experimental and bioinformatic workflow to analyze mutations in cirDNA is established using custom amplicon sequencing. This approach relies on open-software tools to select primers suitable for multiplex PCR using minimal cirDNA as input. In addition, a robust linear model to identify specific genetic alterations from sequencing data of cirDNA was developed. This workflow was applied to design a custom amplicon panel suitable for detection of hotspot mutations relevant for colorectal cancer and analyzed mutations in serial cirDNA samples from a pilot cohort of 34 patients with advanced colorectal cancer. This thesis could show, that in patients with gastrointestinal cancers cirDNA-levels are higher in advanced stages of disease and that cirDNA-levels positively correlate to LDH and CEA in patients with colorectal cancer. However, cirDNA-levels underlie multiple influencing factors and thus impedes to draw individual conclusions. The mutation analysis showed that, recurrent and patient-specific mutational patterns could be detected in most patients.

Furthermore, dynamic changes of mutant allele frequencies in cirDNA were shown to correlate well with disease progression. Finally, sequencing of cirDNA proved to reveal mechanisms of resistance to anti-Epidermal Growth Factor Receptor (EGFR) antibody treatment. Thus, this approach offers a simple and highly customizable method to explore genetic alterations in cirDNA.

6 REFERENCES

1. WHO | Cancer [Internet]. WHO. 2018 Available from: https://www.who.int/cancer/en/ [cited 2019 Oct 20]
2. © Statistisches Bundesamt (Destatis): Statistisches Bundesamt Deutschland - GENESIS-Online [Internet]. 2019 Available from: https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Gesundheit/Todesursachen/Tabellen/sterbefaelle-krebs-insgesamt.html [cited 2019 Oct 20]
3. Kumar V, Abbas AK, Aster JC, Robbins SL: Robbins basic pathology. 9th ed. Philadelphia, PA: Elsevier/Saunders; p. 164
4. Ferlay J, Colombet M, Soerjomataram I, Mathers C, Parkin DM, Piñeros M, Znaor A, Bray F: Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. *Int. J. Cancer* ijc.31937, 2018
5. Robert Koch Institut: Bericht zum Krebsgeschehen in Deutschland 2016. Berlin: Robert Koch-Institut;
6. Arnold M, Sierra MS, Laversanne M, Soerjomataram I, Jemal A, Bray F: Global patterns and trends in colorectal cancer incidence and mortality. *Gut* 66: 683–691, 2017
7. Jasperson KW, Tuohy TM, Neklason DW, Burt RW: Hereditary and Familial Colon Cancer. *Gastroenterology* 138: 2044–2058, 2010
8. Lynch H, Lynch P, Lanspa S, Snyder C, Lynch J, Boland C: Review of the Lynch syndrome: history, molecular genetics, screening, differential diagnosis, and medicolegal ramifications. *Clin. Genet.* 76: 1–18, 2009
9. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D: Global cancer statistics. *CA. Cancer J. Clin.* 61: 69–90, 2011
10. Winawer SJ, Fletcher RH, Miller L, Godlee F, Stolar MH, Mulrow CD, Woolf SH, Glick SN, Ganiats TG, Bond JH, Rosen L, Zapka JG, Olsen SJ, Giardiello FM, Sisk JE, Van Antwerp R, Brown-Davis C, Marciniak DA, Mayer RJ: Colorectal cancer screening: clinical guidelines and rationale. *Gastroenterology* 112: 594–642, 1997
11. Siegel RL, Miller KD, Fedewa SA, Ahnen DJ, Meester RGS, Barzi A, Jemal A: Colorectal cancer statistics, 2017. *CA. Cancer J. Clin.* 67: 177–193, 2017
12. Brenner H, Stock C, Hoffmeister M: Effect of screening sigmoidoscopy and

screening colonoscopy on colorectal cancer incidence and mortality: systematic review and meta-analysis of randomised controlled trials and observational studies. *BMJ* 348: g2467–g2467, 2014

13. Coutinho AK, Rocha Lima CM: Metastatic colorectal cancer: systemic treatment in the new millennium. *Cancer Control* 10: 224–38
14. Prenen H, Vecchione L, Van Cutsem E: Role of targeted agents in metastatic colorectal cancer. *Target. Oncol.* 8: 83–96, 2013
15. Strickler JH, Hurwitz HI: Bevacizumab-Based Therapies in the First-Line Treatment of Metastatic Colorectal Cancer. *Oncologist* 17: 513–524, 2012
16. Lièvre A, Bachet J-B, Boige V, Cayre A, Le Corre D, Buc E, Ychou M, Bouché O, Landi B, Louvet C, André T, Bibeau F, Diebold M-D, Rougier P, Ducreux M, Tomasic G, Emile J-F, Penault-Llorca F, Laurent-Puig P: KRAS Mutations As an Independent Prognostic Factor in Patients With Advanced Colorectal Cancer Treated With Cetuximab. *J. Clin. Oncol.* 26: 374–379, 2008
17. Grasselli J, Elez E, Caratù G, Matito J, Santos C, Macarulla T, Vidal J, Garcia M, Viéitez JM, Paéz D, Falcó E, Lopez Lopez C, Aranda E, Jones F, Sikri V, Nuciforo P, Fasani R, Tabernero J, Montagut C, Azuara D, Dienstmann R, Salazar R, Vivancos A: Concordance of blood- and tumor-based detection of RAS mutations to guide anti-EGFR therapy in metastatic colorectal cancer. *Ann. Oncol.* 28: 1294–1301, 2017
18. De Roock W, Claes B, Bernasconi D, De Schutter J, Biesmans B, Fountzilas G, Kalogeras KT, Kotoula V, Papamichael D, Laurent-Puig P, Penault-Llorca F, Rougier P, Vincenzi B, Santini D, Tonini G, Cappuzzo F, Frattini M, Molinari F, Saletti P, De Dosso S, Martini M, Bardelli A, Siena S, Sartore-Bianchi A, Tabernero J, Macarulla T, Di Fiore F, Gangloff AO, Ciardiello F, Pfeiffer P, Qvortrup C, Hansen TP, Van Cutsem E, Piessevaux H, Lambrechts D, Delorenzi M, Tejpar S: Effects of KRAS, BRAF, NRAS, and PIK3CA mutations on the efficacy of cetuximab plus chemotherapy in chemotherapy-refractory metastatic colorectal cancer: a retrospective consortium analysis. *Lancet Oncol.* 11: 753–762, 2010
19. Holch JW, Ricard I, Stintzing S, Modest DP, Heinemann V: The relevance of primary tumour location in patients with metastatic colorectal cancer: A meta-analysis of first-line clinical trials. *Eur. J. Cancer* 70: 87–98, 2017
20. Cohrs RJ, Martin T, Ghahramani P, Bidaut L, Higgins PJ, Shahzad A:

Translational Medicine definition by the European Society for Translational Medicine. *Eur. J. Mol. Clin. Med.* 2: 86, 2014

21. National Cancer Institute: Definition of personalized medicine - NCI Dictionary of Cancer Terms - National Cancer Institute [Internet]. 2019 Available from: https://www.cancer.gov/publications/dictionaries/cancer-terms/def/personalized-medicine [cited 2019 Oct 20]
22. Jackman RJ, Mayo CW: The adenoma-carcinoma sequence in cancer of the colon. *Surg. Gynecol. Obstet.* 93: 327–30, 1951
23. Fearon ER, Vogelstein B: A genetic model for colorectal tumorigenesis. *Cell* 61: 759–767, 1990
24. Hanahan D, Weinberg RA: The hallmarks of cancer. *Cell* 100: 57–70, 2000
25. Fearon ER: Molecular Genetics of Colorectal Cancer. *Annu. Rev. Pathol. Mech. Dis.* 6: 479–507, 2011
26. Hampel H, Frankel WL, Martin E, Arnold M, Khanduja K, Kuebler P, Nakagawa H, Sotamaa K, Prior TW, Westman J, Panescu J, Fix D, Lockman J, Comeras I, de la Chapelle A: Screening for the Lynch Syndrome (Hereditary Nonpolyposis Colorectal Cancer). *N. Engl. J. Med.* 352: 1851–1860, 2005
27. Knudson AG: Heredity and human cancer. *Am. J. Pathol.* 77: 77–84, 1974
28. Lawrence MS, Stojanov P, Polak P, Kryukov G V., Cibulskis K, Sivachenko A, Carter SL, Stewart C, Mermel CH, Roberts SA, Kiezun A, Hammerman PS, McKenna A, Drier Y, Zou L, Ramos AH, Pugh TJ, Stransky N, Helman E, Kim J, Sougnez C, Ambrogio L, Nickerson E, Shefler E, Cortés ML, Auclair D, Saksena G, Voet D, Noble M, DiCara D, Lin P, Lichtenstein L, Heiman DI, Fennell T, Imielinski M, Hernandez B, Hodis E, Baca S, Dulak AM, Lohr J, Landau D-A, Wu CJ, Melendez-Zajgla J, Hidalgo-Miranda A, Koren A, McCarroll SA, Mora J, Lee RS, Crompton B, Onofrio R, Parkin M, Winckler W, Ardlie K, Gabriel SB, Roberts CWM, Biegel JA, Stegmaier K, Bass AJ, Garraway LA, Meyerson M, Golub TR, Gordenin DA, Sunyaev S, Lander ES, Getz G: Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* 499: 214–218, 2013
29. Cancer Genome Atlas Network: Comprehensive molecular characterization of human colon and rectal cancer. *Nature* 487: 330–337, 2012
30. Yaeger R, Chatila WK, Lipsyc MD, Hechtman JF, Cercek A, Sanchez-Vega F, Jayakumaran G, Middha S, Zehir A, Donoghue MTA, You D, Viale A, Kemeny

N, Segal NH, Stadler ZK, Varghese AM, Kundra R, Gao J, Syed A, Hyman DM, Vakiani E, Rosen N, Taylor BS, Ladanyi M, Berger MF, Solit DB, Shia J, Saltz L, Schultz N: Clinical Sequencing Defines the Genomic Landscape of Metastatic Colorectal Cancer. *Cancer Cell* 33: 125-136.e3, 2018

31. Giannakis M, Mu XJ, Shukla SA, Qian ZR, Cohen O, Nishihara R, Bahl S, Cao Y, Amin-Mansour A, Yamauchi M, Sukawa Y, Stewart C, Rosenberg M, Mima K, Inamura K, Nosho K, Nowak JA, Lawrence MS, Giovannucci EL, Chan AT, Ng K, Meyerhardt JA, Van Allen EM, Getz G, Gabriel SB, Lander ES, Wu CJ, Fuchs CS, Ogino S, Garraway LA: Genomic Correlates of Immune-Cell Infiltrates in Colorectal Carcinoma. *Cell Rep.* 15: 857–865, 2016
32. Nusse R, Clevers H: Wnt/β-Catenin Signaling, Disease, and Emerging Therapeutic Modalities. *Cell* 169: 985–999, 2017
33. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E, Antipin Y, Reva B, Goldberg AP, Sander C, Schultz N: The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data. *Cancer Discov.* 2: 401–404, 2012
34. Kinzler KW, Nilbert MC, Su LK, Vogelstein B, Bryan TM, Levy DB, Smith KJ, Preisinger AC, Hedge P, McKechnie D: Identification of FAP locus genes from chromosome 5q21. *Science* 253: 661–5, 1991
35. Forbes SA, Beare D, Boutselakis H, Bamford S, Bindal N, Tate J, Cole CG, Ward S, Dawson E, Ponting L, Stefancsik R, Harsha B, Kok CY, Jia M, Jubb H, Sondka Z, Thompson S, De T, Campbell PJ: COSMIC: somatic cancer genetics at high-resolution. *Nucleic Acids Res.* 45: D777–D783, 2017
36. Kastenhuber ER, Lowe SW: Putting p53 in Context. *Cell* 170: 1062–1078, 2017
37. Malkin D, Li F, Strong L, Fraumeni J, Nelson C, Kim D, Kassel J, Gryka M, Bischoff F, Tainsky M, Et A: Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. *Science (80-.).* 250: 1233–1238, 1990
38. Guha T, Malkin D: Inherited TP53 Mutations and the Li–Fraumeni Syndrome. *Cold Spring Harb. Perspect. Med.* 7: a026187, 2017
39. Kuismanen SA, Holmberg MT, Salovaara R, de la Chapell A, Peltomäki P: Genetic and epigenetic modification of MLH1 accounts for a major share of microsatellite-unstable colorectal cancers. *Am. J. Pathol.* 156: 1773–1779, 2000
40. Hoeijmakers JHJ: Genome maintenance mechanisms for preventing cancer.

Nature 411: 366–374, 2001

41. Chalmers ZR, Connelly CF, Fabrizio D, Gay L, Ali SM, Ennis R, Schrock A, Campbell B, Shlien A, Chmielecki J, Huang F, He Y, Sun J, Tabori U, Kennedy M, Lieber DS, Roels S, White J, Otto GA, Ross JS, Garraway L, Miller VA, Stephens PJ, Frampton GM: Analysis of 100,000 human cancer genomes reveals the landscape of tumor mutational burden. *Genome Med.* 9: 34, 2017
42. Jeffreys AJ, Wilson V, Thein SL: Hypervariable “minisatellite” regions in human DNA. *Nature* 314: 67–73
43. Alberto Morán, Paloma Ortega, Carmen de Juan, Tamara Fernández-Marcelo, Cristina Frías, Andrés Sánchez-Pernaute, Antonio José Torres, Eduardo Díaz-Rubio, Pilar Iniesta MB: Differential colorectal carcinogenesis: Molecular basis and clinical relevance. *World J. Gastrointest. Oncol.* 2: 151, 2010
44. Tran B, Kopetz S, Tie J, Gibbs P, Jiang Z-Q, Lieu CH, Agarwal A, Maru DM, Sieber O, Desai J: Impact of BRAF mutation and microsatellite instability on the pattern of metastatic spread and prognosis in metastatic colorectal cancer. *Cancer* 117: 4623–4632, 2011
45. Cortes-Ciriano I, Lee S, Park W-Y, Kim T-M, Park PJ: A molecular portrait of microsatellite instability across multiple cancers. *Nat. Commun.* 8: 15180, 2017
46. Watanabe T, Wu T-T, Catalano PJ, Ueki T, Satriano R, Haller DG, Benson AB, Hamilton SR: Molecular Predictors of Survival after Adjuvant Chemotherapy for Colon Cancer. *N. Engl. J. Med.* 344: 1196–1206, 2001
47. Popat S, Hubner R, Houlston RS: Systematic Review of Microsatellite Instability and Colorectal Cancer Prognosis. *J. Clin. Oncol.* 23: 609–618, 2005
48. Klingbiel D, Saridaki Z, Roth AD, Bosman FT, Delorenzi M, Tejpar S: Prognosis of stage II and III colon cancer treated with adjuvant 5-fluorouracil or FOLFIRI in relation to microsatellite status: results of the PETACC-3 trial. *Ann. Oncol.* 26: 126–132, 2015
49. Sargent DJ, Marsoni S, Monges G, Thibodeau SN, Labianca R, Hamilton SR, French AJ, Kabat B, Foster NR, Torri V, Ribic C, Grothey A, Moore M, Zaniboni A, Seitz J-F, Sinicrope F, Gallinger S: Defective Mismatch Repair As a Predictive Marker for Lack of Efficacy of Fluorouracil-Based Adjuvant Therapy in Colon Cancer. *J. Clin. Oncol.* 28: 3219–3226, 2010
50. Müller CI, Schulmann K, Reinacher-Schick A, Andre N, Arnold D, Tannapfel A, Arkenau H, Hahn SA, Schmoll SH-J, Porschen R, Schmiegel W, Graeven U:

Predictive and prognostic value of microsatellite instability in patients with advanced colorectal cancer treated with a fluoropyrimidine and oxaliplatin containing first-line chemotherapy. A report of the AIO Colorectal Study Group. *Int. J. Colorectal Dis.* 23: 1033–1039, 2008

51. Fang JY, Richardson BC: The MAPK signalling pathways and colorectal cancer. *Lancet Oncol.* 6: 322–327, 2005
52. Therkildsen C, Bergmann TK, Henrichsen-Schnack T, Ladelund S, Nilbert M: The predictive value of KRAS, NRAS, BRAF, PIK3CA and PTEN for anti-EGFR treatment in metastatic colorectal cancer: A systematic review and meta-analysis. *Acta Oncol. (Madr).* 53: 852–864, 2014
53. Pietrantonio F, Petrelli F, Coinu A, Di Bartolomeo M, Borgonovo K, Maggi C, Cabiddu M, Iacovelli R, Bossi I, Lonati V, Ghilardi M, de Braud F, Barni S: Predictive role of BRAF mutations in patients with advanced colorectal cancer receiving cetuximab and panitumumab: A meta-analysis. *Eur. J. Cancer* 51: 587–594, 2015
54. Yokota T, Ura T, Shibata N, Takahari D, Shitara K, Nomura M, Kondo C, Mizota A, Utsunomiya S, Muro K, Yatabe Y: BRAF mutation is a powerful prognostic factor in advanced and recurrent colorectal cancer. *Br. J. Cancer* 104: 856–862, 2011
55. Jones JC, Renfro LA, Al-Shamsi HO, Schrock AB, Rankin A, Zhang BY, Kasi PM, Voss JS, Leal AD, Sun J, Ross J, Ali SM, Hubbard JM, Kipp BR, McWilliams RR, Kopetz S, Wolff RA, Grothey A: Non-V600 BRAF Mutations Define a Clinically Distinct Molecular Subtype of Metastatic Colorectal Cancer. *J. Clin. Oncol.* 35: 2624–2630, 2017
56. Silvestris N, Tommasi S, Petriella D, Santini D, Fistola E, Russo A, Numico G, Tonini G, Maiello E, Colucci G: The Dark Side of the Moon: The PI3K/PTEN/AKT Pathway in Colorectal Carcinoma. *Oncology* 77: 69–74, 2009
57. Liaw D, Marsh DJ, Li J, Dahia PL, Wang SI, Zheng Z, Bose S, Call KM, Tsou HC, Peacocke M, Eng C, Parsons R: Germline mutations of the PTEN gene in Cowden disease, an inherited breast and thyroid cancer syndrome. *Nat. Genet.* 16: 64–7, 1997
58. Vidal J, Bellosillo B, Santos Vivas C, García-Alfonso P, Carrato A, Cano MT, García-Carbonero R, Élez E, Losa F, Massutí B, Valladares-Ayerbes M, Viéitez JM, Manzano JL, Azuara D, Gallego J, Pairet S, Capellá G, Salazar R,

Tabernero J, Aranda E, Montagut C: Ultra-selection of metastatic colorectal cancer patients using next-generation sequencing to improve clinical efficacy of anti-EGFR therapy. *Ann. Oncol.* 30: 439–446, 2019

59. Mohamed A, Twardy B, AbdAllah N, Akhras A, Ismail H, Zordok M, Schrapp K, Attumi T, Tesfaye A, El-Rayes B: Clinical Impact of PI3K/BRAF Mutations in RAS Wild Metastatic Colorectal Cancer: Meta-analysis Results. *J. Gastrointest. Cancer* 2018
60. Bierie B, Moses HL: TGFβ: the molecular Jekyll and Hyde of cancer. *Nat. Rev. Cancer* 6: 506–520, 2006
61. Labelle M, Begum S, Hynes RO: Direct Signaling between Platelets and Cancer Cells Induces an Epithelial-Mesenchymal-Like Transition and Promotes Metastasis. *Cancer Cell* 20: 576–590, 2011
62. Xu Y, Pasche B: TGF-β signaling alterations and susceptibility to colorectal cancer. *Hum. Mol. Genet.* 16: R14–R20, 2007
63. Brosens LA: Juvenile polyposis syndrome. *World J. Gastroenterol.* 17: 4839, 2011
64. Zhang B, Zhang B, Chen X, Bae S, Singh K, Washington MK, Datta PK: Loss of Smad4 in colorectal cancer induces resistance to 5-fluorouracil through activating Akt pathway. *Br. J. Cancer* 110: 946–957, 2014
65. Cao J, Ge M-H, Ling Z-Q: Fbxw7 Tumor Suppressor. *Medicine (Baltimore).* 95: e2496, 2016
66. Yamada M, Sekine S, Ogawa R, Taniguchi H, Kushima R, Tsuda H, Kanai Y: Frequent activating GNAS mutations in villous adenoma of the colorectum. *J. Pathol.* 228: n/a-n/a, 2012
67. Wilson CH, McIntyre RE, Arends MJ, Adams DJ: The activating mutation R201C in GNAS promotes intestinal tumourigenesis in ApcMin/+ mice through activation of Wnt and ERK1/2 MAPK pathways. *Oncogene* 29: 4567–4575, 2010
68. Leitlinienprogramm Onkologie (Deutsche Krebsgesellschaft, Deutsche Krebshilfe, AWMF): S3-Leitlinie Kolorektales Karzinom, Langversion 2.1, 2019, AWMF Registrierungsnummer: 021/007OL, http://www.leitlinienprogramm-onkologie.de/leitlinien/kolorektales-karzinom/ [cited 2019 Oct 20]
69. Longley DB, Harkin DP, Johnston PG: 5-Fluorouracil: mechanisms of action and clinical strategies. *Nat. Rev. Cancer* 3: 330–338, 2003
70. Jennings BA, Kwok CS, Willis G, Matthews V, Wawruch P, Loke YK: Functional

polymorphisms of folate metabolism and response to chemotherapy for colorectal cancer, a systematic review and meta-analysis. *Pharmacogenet. Genomics* 22: 290–304, 2012

71. Pullarkat ST, Stoehlmacher J, Ghaderi V, Xiong Y-P, Ingles S a, Sherrod A, Warren R, Tsao-Wei D, Groshen S, Lenz H-J: Thymidylate synthase gene polymorphism determines response and toxicity of 5-FU chemotherapy. *Pharmacogenomics J.* 1: 65–70, 2001
72. Salonga D, Danenberg KD, Johnson M, Metzger R, Groshen S, Tsao-Wei DD, Lenz HJ, Leichman CG, Leichman L, Diasio RB, Danenberg P V: Colorectal tumors responding to 5-fluorouracil have low gene expression levels of dihydropyrimidine dehydrogenase, thymidylate synthase, and thymidine phosphorylase. *Clin. Cancer Res.* 6: 1322–7, 2000
73. Ooyama A, Okayama Y, Takechi T, Sugimoto Y, Oka T, Fukushima M: Genome-wide screening of loci associated with drug resistance to 5-fluorouracil-based drugs. *Cancer Sci.* 98: 577–583, 2007
74. Scartozzi M, Maccaroni E, Giampieri R, Pistelli M, Bittoni A, Del Prete M, Berardi R, Cascinu S: 5-fluorouracil pharmacogenomics: still rocking after all these years? *Pharmacogenomics* 12: 251–265, 2011
75. Mandel P, Metais P: Les acides nucléiques du plasma sanguin chez l'homme. *C. R. Seances Soc. Biol. Fil.* 142: 241–3, 1948
76. Leon SA, Shapiro B, Sklaroff DM, Yaros MJ: Free DNA in the serum of cancer patients and the effect of therapy. *Cancer Res.* 37: 646–50, 1977
77. Wan JCM, Massie C, Garcia-Corbacho J, Mouliere F, Brenton JD, Caldas C, Pacey S, Baird R, Rosenfeld N: Liquid biopsies come of age: towards implementation of circulating tumour DNA. *Nat. Rev. Cancer* 17: 223–238, 2017
78. 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* 46: 318–322, 1989
79. Jahr S, Hentze H, Englisch S, Hardt D, Fackelmayer FO, Hesch RD, Knippers R: DNA fragments in the blood plasma of cancer patients: quantitations and evidence for their origin from apoptotic and necrotic cells. *Cancer Res.* 61: 1659–65, 2001
80. Thakur BK, Zhang H, Becker A, Matei I, Huang Y, Costa-Silva B, Zheng Y, Hoshino A, Brazier H, Xiang J, Williams C, Rodriguez-Barrueco R, Silva JM,

Zhang W, Hearn S, Elemento O, Paknejad N, Manova-Todorova K, Welte K, Bromberg J, Peinado H, Lyden D: Double-stranded DNA in exosomes: a novel biomarker in cancer detection. *Cell Res.* 24: 766–769, 2014

81. Stroun M, Lyautey J, Lederrey C, Olson-Sand A, Anker P: About the possible origin and mechanism of circulating DNA. *Clin. Chim. Acta* 313: 139–142, 2001
82. Snyder MW, Kircher M, Hill AJ, Daza RM, Shendure J: Cell-free DNA Comprises an In Vivo Nucleosome Footprint that Informs Its Tissues-Of-Origin. *Cell* 164: 57–68, 2016
83. Sun K, Jiang P, Chan KCA, Wong J, Cheng YKY, Liang RHS, Chan W, Ma ESK, Chan SL, Cheng SH, Chan RWY, Tong YK, Ng SSM, Wong RSM, Hui DSC, Leung TN, Leung TY, Lai PBS, Chiu RWK, Lo YMD: Plasma DNA tissue mapping by genome-wide methylation sequencing for noninvasive prenatal, cancer, and transplantation assessments. *Proc. Natl. Acad. Sci.* 112: E5503–E5512, 2015
84. Chelobanov BP, Laktionov PP, Vlasov V V.: Proteins involved in binding and cellular uptake of nucleic acids. *Biochem.* 71: 583–596, 2006
85. Rykova EY, Morozkin ES, Ponomaryova AA, Loseva EM, Zaporozhchenko IA, Cherdyntseva N V, Vlassov V V, Laktionov PP: Cell-free and cell-bound circulating nucleic acid complexes: mechanisms of generation, concentration and content. *Expert Opin. Biol. Ther.* 12: S141–S153, 2012
86. Jiang P, Chan CWM, Chan KCA, Cheng SH, Wong J, Wong VW-S, Wong GLH, Chan SL, Mok TSK, Chan HLY, Lai PBS, Chiu RWK, Lo YMD: Lengthening and shortening of plasma DNA in hepatocellular carcinoma patients. *Proc. Natl. Acad. Sci.* 112: E1317–E1325, 2015
87. Underhill HR, Kitzman JO, Hellwig S, Welker NC, Daza R, Baker DN, Gligorich KM, Rostomily RC, Bronner MP, Shendure J: Fragment Length of Circulating Tumor DNA. *PLOS Genet.* 12: e1006162, 2016
88. Thierry AR, El Messaoudi S, Gahan PB, Anker P, Stroun M: Origins, structures, and functions of circulating DNA in oncology. *Cancer Metastasis Rev.* 35: 347–376, 2016
89. Nishimoto S, Fukuda D, Higashikuni Y, Tanaka K, Hirata Y, Murata C, Kim-Kaneyama J, Sato F, Bando M, Yagi S, Soeki T, Hayashi T, Imoto I, Sakaue H, Shimabukuro M, Sata M: Obesity-induced DNA released from adipocytes stimulates chronic adipose tissue inflammation and insulin resistance. *Sci. Adv.*

2: e1501332, 2016

90. Tsokos GC: Systemic Lupus Erythematosus. *N. Engl. J. Med.* 365: 2110–2121, 2011
91. Garcia-Olmo DC, Dominguez C, Garcia-Arranz M, Anker P, Stroun M, Garcia-Verdugo JM, Garcia-Olmo D: Cell-Free Nucleic Acids Circulating in the Plasma of Colorectal Cancer Patients Induce the Oncogenic Transformation of Susceptible Cultured Cells. *Cancer Res.* 70: 560–567, 2010
92. Bettegowda C, Sausen M, Leary RJ, Kinde I, Wang Y, Agrawal N, Bartlett BR, Wang H, Luber B, Alani RM, Antonarakis ES, Azad NS, Bardelli A, Brem H, Cameron JL, Lee CC, Fecher LA, Gallia GL, Gibbs P, Le D, Giuntoli RL, Goggins M, Hogarty MD, Holdhoff M, Hong S-M, Jiao Y, Juhl HH, Kim JJ, Siravegna G, Laheru DA, Lauricella C, Lim M, Lipson EJ, Marie SKN, Netto GJ, Oliner KS, Olivi A, Olsson L, Riggins GJ, Sartore-Bianchi A, Schmidt K, Shih L-M, Oba-Shinjo SM, Siena S, Theodorescu D, Tie J, Harkins TT, Veronese S, Wang T-L, Weingart JD, Wolfgang CL, Wood LD, Xing D, Hruban RH, Wu J, Allen PJ, Schmidt CM, Choti MA, Velculescu VE, Kinzler KW, Vogelstein B, Papadopoulos N, Diaz LA: Detection of Circulating Tumor DNA in Early- and Late-Stage Human Malignancies. *Sci. Transl. Med.* 6: 224ra24, 2014
93. Mouliere F, El Messaoudi S, Pang D, Dritschilo A, Thierry AR: Multi-marker analysis of circulating cell-free DNA toward personalized medicine for colorectal cancer. *Mol. Oncol.* 8: 927–941, 2014
94. Diehl F, Li M, Dressman D, He Y, Shen D, Szabo S, Diaz L a, Goodman SN, David K a, Juhl H, Kinzler KW, Vogelstein B: Detection and quantification of mutations in the plasma of patients with colorectal tumors. *Proc. Natl. Acad. Sci.* 102: 16368–16373, 2005
95. Phallen J, Sausen M, Adleff V, Leal A, Hruban C, White J, Anagnostou V, Fiksel J, Cristiano S, Papp E, Speir S, Reinert T, Orntoft M-BW, Woodward BD, Murphy D, Parpart-Li S, Riley D, Nesselbush M, Sengamalay N, Georgiadis A, Li QK, Madsen MR, Mortensen FV, Huiskens J, Punt C, van Grieken N, Fijneman R, Meijer G, Husain H, Scharpf RB, Diaz LA, Jones S, Angiuoli S, Ørntoft T, Nielsen HJ, Andersen CL, Velculescu VE: Direct detection of early-stage cancers using circulating tumor DNA. *Sci. Transl. Med.* 9: eaan2415, 2017
96. Murtaza M, Dawson S-JJ, Tsui DWYY, Gale D, Forshew T, Piskorz AM, Parkinson C, Chin S-FF, Kingsbury Z, Wong ASCC, Marass F, Humphray S,

Hadfield J, Bentley D, Chin TM, Brenton JD, Caldas C, Rosenfeld N: Non-invasive analysis of acquired resistance to cancer therapy by sequencing of plasma DNA. *Nature* 497: 108–112, 2013

97. Misale S, Yaeger R, Hobor S, Scala E, Janakiraman M, Liska D, Valtorta E, Schiavo R, Buscarino M, Siravegna G, Bencardino K, Cercek A, Chen C-T, Veronese S, Zanon C, Sartore-Bianchi A, Gambacorta M, Gallicchio M, Vakiani E, Boscaro V, Medico E, Weiser M, Siena S, Di Nicolantonio F, Solit D, Bardelli A: Emergence of KRAS mutations and acquired resistance to anti-EGFR therapy in colorectal cancer. *Nature* 486: 532–536, 2012
98. Lo YMD, Zhang J, Leung TN, Lau TK, Chang AMZ, Hjelm NM: Rapid Clearance of Fetal DNA from Maternal Plasma. *Am. J. Hum. Genet.* 64: 218–224, 1999
99. Diehl F, Schmidt K, Choti M a, Romans K, Goodman S, Li M, Thornton K, Agrawal N, Sokoll L, Szabo S a, Kinzler KW, Vogelstein B, Diaz Jr LA: Circulating mutant DNA to assess tumor dynamics. *Nat. Med.* 14: 985–990, 2008
100. Parkinson CA, Gale D, Piskorz AM, Biggs H, Hodgkin C, Addley H, Freeman S, Moyle P, Sala E, Sayal K, Hosking K, Gounaris I, Jimenez-Linan M, Earl HM, Qian W, Rosenfeld N, Brenton JD: Exploratory Analysis of TP53 Mutations in Circulating Tumour DNA as Biomarkers of Treatment Response for Patients with Relapsed High-Grade Serous Ovarian Carcinoma: A Retrospective Study. *PLOS Med.* 13: e1002198, 2016
101. Diaz LA, Bardelli A: Liquid Biopsies: Genotyping Circulating Tumor DNA. *J. Clin. Oncol.* 32: 579–586, 2014
102. Xi L, Pham TH-T, Payabyab EC, Sherry RM, Rosenberg SA, Raffeld M: Circulating Tumor DNA as an Early Indicator of Response to T-cell Transfer Immunotherapy in Metastatic Melanoma. *Clin. Cancer Res.* 22: 5480–5486, 2016
103. Tie J, Kinde I, Wang Y, Wong HL, Roebert J, Christie M, Tacey M, Wong R, Singh M, Karapetis CS, Desai J, Tran B, Strausberg RL, Diaz LA, Papadopoulos N, Kinzler KW, Vogelstein B, Gibbs P: Circulating tumor DNA as an early marker of therapeutic response in patients with metastatic colorectal cancer. *Ann. Oncol.* 26: 1715–1722, 2015
104. Tóth K, Patai Á V, Kalmár A, Barták BK, Nagy ZB, Galamb O, Wichmann B, Tulassay Z, Molnár B: Circadian Rhythm of Methylated Septin 9, Cell-Free DNA Amount and Tumor Markers in Colorectal Cancer Patients. *Pathol. Oncol. Res.*

23: 699–706, 2017

105. Jung M, Klotzek S, Lewandowski M, Fleischhacker M, Jung K: Changes in Concentration of DNA in Serum and Plasma during Storage of Blood Samples. *Clin. Chem.* 49: 1028–1029, 2003
106. Lam NYL, Rainer TH, Chiu RWK, Lo YMD: EDTA Is a Better Anticoagulant than Heparin or Citrate for Delayed Blood Processing for Plasma DNA Analysis. *Clin. Chem.* 50: 256–257, 2004
107. Norton SE, Lechner JM, Williams T, Fernando MR: A stabilizing reagent prevents cell-free DNA contamination by cellular DNA in plasma during blood sample storage and shipping as determined by digital PCR. *Clin. Biochem.* 46: 1561–1565, 2013
108. El Messaoudi S, Rolet F, Mouliere F, Thierry AR: Circulating cell free DNA: Preanalytical considerations. *Clin. Chim. Acta* 424: 222–230, 2013
109. Page K, Guttery DS, Zahra N, Primrose L, Elshaw SR, Pringle JH, Blighe K, Marchese SD, Hills A, Woodley L, Stebbing J, Coombes RC, Shaw JA: Influence of Plasma Processing on Recovery and Analysis of Circulating Nucleic Acids. *PLoS One* 8: e77963, 2013
110. Warton K, Graham L-J, Yuwono N, Samimi G: Comparison of 4 commercial kits for the extraction of circulating DNA from plasma. *Cancer Genet.* 228–229: 143–150, 2018
111. Diefenbach RJ, Lee JH, Kefford RF, Rizos H: Evaluation of commercial kits for purification of circulating free DNA. *Cancer Genet.* 228–229: 21–27, 2018
112. Robin JD, Ludlow AT, LaRanger R, Wright WE, Shay JW: Comparison of DNA Quantification Methods for Next Generation Sequencing. *Sci. Rep.* 6: 24067, 2016
113. Kloten V, Rüchel N, Brüchle NO, Gasthaus J, Freudenmacher N, Steib F, Mijnes J, Eschenbruch J, Binnebösel M, Knüchel R, Dahl E: Liquid biopsy in colon cancer: comparison of different circulating DNA extraction systems following absolute quantification of KRAS mutations using Intplex allele-specific PCR. *Oncotarget* 8: 86253, 2017
114. Alix-Panabières C, Pantel K: Characterization of single circulating tumor cells. *FEBS Lett.* 591: 2241–2250, 2017
115. Best MG, Sol N, Kooi I, Tannous J, Westerman BA, Rustenburg F, Schellen P, Verschueren H, Post E, Koster J, Ylstra B, Ameziane N, Dorsman J, Smit EF,

Verheul HM, Noske DP, Reijneveld JC, Nilsson RJA, Tannous BA, Wesseling P, Wurdinger T: RNA-Seq of Tumor-Educated Platelets Enables Blood-Based Pan-Cancer, Multiclass, and Molecular Pathway Cancer Diagnostics. *Cancer Cell* 28: 666–676, 2015

116. Shen SY, Singhania R, Fehringer G, Chakravarthy A, Roehrl MHA, Chadwick D, Zuzarte PC, Borgida A, Wang TT, Li T, Kis O, Zhao Z, Spreafico A, Medina T da S, Wang Y, Roulois D, Ettayebi I, Chen Z, Chow S, Murphy T, Arruda A, O'Kane GM, Liu J, Mansour M, McPherson JD, O'Brien C, Leighl N, Bedard PL, Fleshner N, Liu G, Minden MD, Gallinger S, Goldenberg A, Pugh TJ, Hoffman MM, Bratman S V., Hung RJ, De Carvalho DD: Sensitive tumour detection and classification using plasma cell-free DNA methylomes. *Nature* 563: 579–583, 2018
117. Lehmann-Werman R, Neiman D, Zemmour H, Moss J, Magenheim J, Vaknin-Dembinsky A, Rubertsson S, Nellgård B, Blennow K, Zetterberg H, Spalding K, Haller MJ, Wasserfall CH, Schatz DA, Greenbaum CJ, Dorrell C, Grompe M, Zick A, Hubert A, Maoz M, Fendrich V, Bartsch DK, Golan T, Ben Sasson SA, Zamir G, Razin A, Cedar H, Shapiro AMJ, Glaser B, Shemer R, Dor Y: Identification of tissue-specific cell death using methylation patterns of circulating DNA. *Proc. Natl. Acad. Sci.* 113: E1826–E1834, 2016
118. Chevillet JR, Kang Q, Ruf IK, Briggs HA, Vojtech LN, Hughes SM, Cheng HH, Arroyo JD, Meredith EK, Gallichotte EN, Pogosova-Agadjanyan EL, Morrissey C, Stirewalt DL, Hladik F, Yu EY, Higano CS, Tewari M: Quantitative and stoichiometric analysis of the microRNA content of exosomes. *Proc. Natl. Acad. Sci.* 111: 14888–14893, 2014
119. Schwarzenbach H, Hoon DSB, Pantel K: Cell-free nucleic acids as biomarkers in cancer patients. *Nat. Rev. Cancer* 11: 426–437, 2011
120. Spindler KLG, Pallisgaard N, Andersen RF, Brandslund I, Jakobsen A: Circulating Free DNA as Biomarker and Source for Mutation Detection in Metastatic Colorectal Cancer. *PLoS One* 10: e0108247, 2015
121. El Messaoudi S, Mouliere F, Du Manoir S, Bascoul-Mollevi C, Gillet B, Nouaille M, Fiess C, Crapez E, Bibeau F, Theillet C, Mazard T, Pezet D, Mathonnet M, Ychou M, Thierry AR: Circulating DNA as a Strong Multimarker Prognostic Tool for Metastatic Colorectal Cancer Patient Management Care. *Clin. Cancer Res.* 22: 3067–3077, 2016

122. Garlan F, Laurent-Puig P, Sefrioui D, Siauve N, Didelot A, Sarafan-Vasseur N, Michel P, Perkins G, Mulot C, Blons H, Taieb J, Di Fiore F, Taly V, Zaanan A: Early Evaluation of Circulating Tumor DNA as Marker of Therapeutic Efficacy in Metastatic Colorectal Cancer Patients (PLACOL Study). *Clin. Cancer Res.* 23: 5416–5425, 2017
123. Lecomte T, Berger A, Zinzindohoué F, Micard S, Landi B, Blons H, Beaune P, Cugnenc P-H, Laurent-Puig P: Detection of free-circulating tumor-associated DNA in plasma of colorectal cancer patients and its association with prognosis. *Int. J. Cancer* 100: 542–548, 2002
124. Gray ES, Rizos H, Reid AL, Boyd SC, Pereira MR, Lo J, Tembe V, Freeman J, Lee JHJ, Scolyer RA, Siew K, Lomma C, Cooper A, Khattak MA, Meniawy TM, Long G V., Carlino MS, Millward M, Ziman M: Circulating tumor DNA to monitor treatment response and detect acquired resistance in patients with metastatic melanoma. *Oncotarget* 6: 42008–42018, 2015
125. Forshew T, Murtaza M, Parkinson C, Gale D, Tsui DWY, Kaper F, Dawson S-J, Piskorz a. M, Jimenez-Linan M, Bentley D, Hadfield J, May a. P, Caldas C, Brenton JD, Rosenfeld N: Noninvasive Identification and Monitoring of Cancer Mutations by Targeted Deep Sequencing of Plasma DNA. *Sci. Transl. Med.* 4: 136ra68-136ra68, 2012
126. Newman AM, Bratman S V, To J, Wynne JF, Eclov NCW, Modlin L a, Liu CL, Neal JW, Wakelee H a, Merritt RE, Shrager JB, Loo BW, Alizadeh A a, Diehn M: An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage. *Nat. Med.* 20: 548–554, 2014
127. Dawson S-J, Tsui DWY, Murtaza M, Biggs H, Rueda OM, Chin S-F, Dunning MJ, Gale D, Forshew T, Mahler-Araujo B, Rajan S, Humphray S, Becq J, Halsall D, Wallis M, Bentley D, Caldas C, Rosenfeld N: Analysis of Circulating Tumor DNA to Monitor Metastatic Breast Cancer. *N. Engl. J. Med.* 368: 1199–1209, 2013
128. Jeffery M, Hickey BE, Hider PN, See AM: Follow-up strategies for patients treated for non-metastatic colorectal cancer. *Cochrane Database Syst. Rev.* 11: CD002200, 2016
129. Tie J, Wang Y, Tomasetti C, Li L, Springer S, Kinde I, Silliman N, Tacey M, Wong H-L, Christie M, Kosmider S, Skinner I, Wong R, Steel M, Tran B, Desai J, Jones I, Haydon A, Hayes T, Price TJ, Strausberg RL, Diaz LA, Papadopoulos N,

Kinzler KW, Vogelstein B, Gibbs P: Circulating tumor DNA analysis detects minimal residual disease and predicts recurrence in patients with stage II colon cancer. *Sci. Transl. Med.* 8: 346ra92, 2016

130. Garcia-Murillas I, Schiavon G, Weigelt B, Ng C, Hrebien S, Cutts RJ, Cheang M, Osin P, Nerurkar A, Kozarewa I, Garrido JA, Dowsett M, Reis-Filho JS, Smith IE, Turner NC: Mutation tracking in circulating tumor DNA predicts relapse in early breast cancer. *Sci. Transl. Med.* 7: 302ra133, 2015
131. Reinert T, Schøler L V, Thomsen R, Tobiasen H, Vang S, Nordentoft I, Lamy P, Kannerup A-S, Mortensen F V, Stribolt K, Hamilton-Dutoit S, Nielsen HJ, Laurberg S, Pallisgaard N, Pedersen JS, Ørntoft TF, Andersen CL: Analysis of circulating tumour DNA to monitor disease burden following colorectal cancer surgery. *Gut* 65: 625–634, 2016
132. Gormally E, Vineis P, Matullo G, Veglia F, Caboux E, Le Roux E, Peluso M, Garte S, Guarrera S, Munnia A, Airoldi L, Autrup H, Malaveille C, Dunning A, Overvad K, Tjønneland A, Lund E, Clavel-Chapelon F, Boeing H, Trichopoulou A, Palli D, Krogh V, Tumino R, Panico S, Bueno-de-Mesquita HB, Peeters PH, Pera G, Martinez C, Dorronsoro M, Barricarte A, Navarro C, Quirós JR, Hallmans G, Day NE, Key TJ, Saracci R, Kaaks R, Riboli E, Hainaut P: TP53 and KRAS2 Mutations in Plasma DNA of Healthy Subjects and Subsequent Cancer Occurrence: A Prospective Study. *Cancer Res.* 66: 6871–6876, 2006
133. Cohen JD, Li L, Wang Y, Thoburn C, Afsari B, Danilova L, Douville C, Javed AA, Wong F, Mattox A, Hruban RH, Wolfgang CL, Goggins MG, Dal Molin M, Wang T-L, Roden R, Klein AP, Ptak J, Dobbyn L, Schaefer J, Silliman N, Popoli M, Vogelstein JT, Browne JD, Schoen RE, Brand RE, Tie J, Gibbs P, Wong H-L, Mansfield AS, Jen J, Hanash SM, Falconi M, Allen PJ, Zhou S, Bettegowda C, Diaz LA, Tomasetti C, Kinzler KW, Vogelstein B, Lennon AM, Papadopoulos N: Detection and localization of surgically resectable cancers with a multi-analyte blood test. *Science (80-.).* 359: 926–930, 2018
134. Churko JM, Mantalas GL, Snyder MP, Wu JC: Overview of High Throughput Sequencing Technologies to Elucidate Molecular Pathways in Cardiovascular Diseases. *Circ. Res.* 112: 1613–1623, 2013
135. Morey M, Fernández-Marmiesse A, Castiñeiras D, Fraga JM, Couce ML, Cocho JA: A glimpse into past, present, and future DNA sequencing. *Mol. Genet. Metab.* 110: 3–24, 2013

136. Hindson BJ, Ness KD, Masquelier DA, Belgrader P, Heredia NJ, Makarewicz AJ, Bright IJ, Lucero MY, Hiddessen AL, Legler TC, Kitano TK, Hodel MR, Petersen JF, Wyatt PW, Steenblock ER, Shah PH, Bousse LJ, Troup CB, Mellen JC, Wittmann DK, Erndt NG, Cauley TH, Koehler RT, So AP, Dube S, Rose KA, Montesclaros L, Wang S, Stumbo DP, Hodges SP, Romine S, Milanovich FP, White HE, Regan JF, Karlin-Neumann GA, Hindson CM, Saxonov S, Colston BW: High-Throughput Droplet Digital PCR System for Absolute Quantitation of DNA Copy Number. *Anal. Chem.* 83: 8604–8610, 2011
137. Sysmex Deutschland GmbH: ONCOBEAM RAS CRC IVD KIT [Internet]. 2019 Available from: https://www.sysmex.de/produkte/oncobeamTM-ras-crc-ivd-kit-4063.html [cited 2019 Oct 20]
138. Simon R, Roychowdhury S: Implementing personalized cancer genomics in clinical trials. *Nat. Rev. Drug Discov.* 12: 358–369, 2013
139. De Mattos-Arruda L, Weigelt B, Cortes J, Won HH, Ng CKY, Nuciforo P, Bidard F-C, Aura C, Saura C, Peg V, Piscuoglio S, Oliveira M, Smolders Y, Patel P, Norton L, Tabernero J, Berger MF, Seoane J, Reis-Filho JS: Capturing intra-tumor genetic heterogeneity by de novo mutation profiling of circulating cell-free tumor DNA: a proof-of-principle. *Ann. Oncol.* 25: 1729–1735, 2014
140. Murtaza M, Dawson S-J, Pogrebniak K, Rueda OM, Provenzano E, Grant J, Chin S-F, Tsui DWY, Marass F, Gale D, Ali HR, Shah P, Contente-Cuomo T, Farahani H, Shumansky K, Kingsbury Z, Humphray S, Bentley D, Shah SP, Wallis M, Rosenfeld N, Caldas C: Multifocal clonal evolution characterized using circulating tumour DNA in a case of metastatic breast cancer. *Nat. Commun.* 6: 8760, 2015
141. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, Dancey J, Arbuck S, Gwyther S, Mooney M, Rubinstein L, Shankar L, Dodd L, Kaplan R, Lacombe D, Verweij J: New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). *Eur. J. Cancer* 45: 228–247, 2009
142. Forbes SA, Beare D, Gunasekaran P, Leung K, Bindal N, Boutselakis H, Ding M, Bamford S, Cole C, Ward S, Kok CY, Jia M, De T, Teague JW, Stratton MR, McDermott U, Campbell PJ: COSMIC: exploring the world's knowledge of somatic mutations in human cancer. *Nucleic Acids Res.* 43: D805–D811, 2015
143. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen

SG: Primer3—new capabilities and interfaces. *Nucleic Acids Res.* 40: e115–e115, 2012

144. Kaplinski L, Remm M: MultiPLX: Automatic Grouping and Evaluation of PCR Primers. In: *Methods in Molecular Biology*, pp 127–142, 2015
145. Potapov V, Ong JL: Examining Sources of Error in PCR by Single-Molecule Sequencing. *PLoS One* 12: e0169774, 2017
146. Ståhlberg A, Krzyzanowski PM, Jackson JB, Egyud M, Stein L, Godfrey TE: Simple, multiplexed, PCR-based barcoding of DNA enables sensitive mutation detection in liquid biopsies using sequencing. *Nucleic Acids Res.* 44: e105–e105, 2016
147. Bronner IF, Quail MA, Swerdlow H, Turner DJ: Improved Protocols for the Illumina Genome Analyzer Sequencing System. In: *Current Protocols in Human Genetics*, p 2, 2009
148. Herrmann S, Zhan T, Betge J, Rauscher B, Belle S, Gutting T, Schulte N, Jesenofsky R, Härtel N, Gaiser T, Hofheinz R, Ebert MP, Boutros M: Detection of mutational patterns in cell-free DNA of colorectal cancer by custom amplicon sequencing. *Mol. Oncol.* 1878–0261.12539, 2019
149. Andrew S: A quality control tool for high throughput sequence data. *http://Www.Bioinformatics.Babraham.Ac.Uk/Projects/Fastqc/* 2010
150. Bolger AM, Lohse M, Usadel B: Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30: 2114–2120, 2014
151. Li H, Durbin R: Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25: 1754–1760, 2009
152. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R: The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25: 2078–2079, 2009
153. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: Basic local alignment search tool. *J. Mol. Biol.* 215: 403–410, 1990
154. Hulsen T, de Vlieg J, Alkema W: BioVenn – a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams. *BMC Genomics* 9: 488, 2008
155. Steinhart Z, Pavlovic Z, Chandrashekhar M, Hart T, Wang X, Zhang X, Robitaille M, Brown KR, Jaksani S, Overmeer R, Boj SF, Adams J, Pan J, Clevers H, Sidhu S, Moffat J, Angers S: Genome-wide CRISPR screens reveal a Wnt–FZD5

signaling circuit as a druggable vulnerability of RNF43-mutant pancreatic tumors. *Nat. Med.* 23: 60–68, 2017

156. Khagi Y, Kurzrock R, Patel SP: Next generation predictive biomarkers for immune checkpoint inhibition. *Cancer Metastasis Rev.* 36: 179–190, 2017
157. Han S-W, Kim H-P, Shin J-Y, Jeong E-G, Lee W-C, Lee K-H, Won J-K, Kim T-Y, Oh D-Y, Im S-A, Bang Y-J, Jeong S-Y, Park KJ, Park J-G, Kang GH, Seo J-S, Kim J-I, Kim T-Y: Targeted Sequencing of Cancer-Related Genes in Colorectal Cancer Using Next-Generation Sequencing. *PLoS One* 8: e64271, 2013
158. Tougeron D, Lecomte T, Pages JC, Villalva C, Collin C, Ferru A, Tourani JM, Silvain C, Levillain P, Karayan-Tapon L: Effect of low-frequency KRAS mutations on the response to anti-EGFR therapy in metastatic colorectal cancer. *Ann. Oncol.* 24: 1267–1273, 2013
159. Spencer DH, Sehn JK, Abel HJ, Watson MA, Pfeifer JD, Duncavage EJ: Comparison of Clinical Targeted Next-Generation Sequence Data from Formalin-Fixed and Fresh-Frozen Tissue Specimens. *J. Mol. Diagnostics* 15: 623–633, 2013
160. Mardis E, McCombie WR: Agarose Gel Size Selection for DNA Sequencing Libraries. *Cold Spring Harb. Protoc.* 2017: pdb.prot094698, 2017
161. QIAGEN: DNA Clean Up - QIAGEN Online Shop [Internet]. Available from: https://www.qiagen.com/us/shop/sample-technologies/dna/dna-clean-up/ [cited 2019 Oct 20]
162. Kinde I, Wu J, Papadopoulos N, Kinzler KW, Vogelstein B: Detection and quantification of rare mutations with massively parallel sequencing. *Proc. Natl. Acad. Sci.* 108: 9530–9535, 2011
163. Le DT, Uram JN, Wang H, Bartlett BR, Kemberling H, Eyring AD, Skora AD, Luber BS, Azad NS, Laheru D, Biedrzycki B, Donehower RC, Zaheer A, Fisher GA, Crocenzi TS, Lee JJ, Duffy SM, Goldberg RM, de la Chapelle A, Koshiji M, Bhaijee F, Huebner T, Hruban RH, Wood LD, Cuka N, Pardoll DM, Papadopoulos N, Kinzler KW, Zhou S, Cornish TC, Taube JM, Anders RA, Eshleman JR, Vogelstein B, Diaz LA: PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. *N. Engl. J. Med.* 372: 2509–20, 2015
164. Trigg RM, Martinson LJ, Parpart-Li S, Shaw JA: Factors that influence quality and yield of circulating-free DNA: A systematic review of the methodology

literature. *Heliyon* 4: e00699, 2018

165. O'Rawe J, Jiang T, Sun G, Wu Y, Wang W, Hu J, Bodily P, Tian L, Hakonarson H, Johnson WE, Wei Z, Wang K, Lyon GJ: Low concordance of multiple variant-calling pipelines: practical implications for exome and genome sequencing. *Genome Med.* 5: 28, 2013
166. Newman AM, Lovejoy AF, Klass DM, Kurtz DM, Chabon JJ, Scherer F, Stehr H, Liu CL, Bratman S V, Say C, Zhou L, Carter JN, West RB, Sledge Jr GW, Shrager JB, Loo BW, Neal JW, Wakelee HA, Diehn M, Alizadeh AA: Integrated digital error suppression for improved detection of circulating tumor DNA. *Nat. Biotechnol.* 34: 547–555, 2016
167. Hu Y, Ulrich BC, Supplee J, Kuang Y, Lizotte PH, Feeney NB, Guibert NM, Awad MM, Wong K-K, Jänne PA, Paweletz CP, Oxnard GR: False-Positive Plasma Genotyping Due to Clonal Hematopoiesis. *Clin. Cancer Res.* 24: 4437–4443, 2018
168. Chang MT, Bhattarai TS, Schram AM, Bielski CM, Donoghue MTA, Jonsson P, Chakravarty D, Phillips S, Kandoth C, Penson A, Gorelick A, Shamu T, Patel S, Harris C, Gao J, Sumer SO, Kundra R, Razavi P, Li BT, Reales DN, Socci ND, Jayakumaran G, Zehir A, Benayed R, Arcila ME, Chandarlapaty S, Ladanyi M, Schultz N, Baselga J, Berger MF, Rosen N, Solit DB, Hyman DM, Taylor BS: Accelerating Discovery of Functional Mutant Alleles in Cancer. *Cancer Discov.* 8: 174–183, 2018
169. Mouliere F, Chandrananda D, Piskorz AM, Moore EK, Morris J, Ahlborn LB, Mair R, Goranova T, Marass F, Heider K, Wan JCM, Supernat A, Hudecova I, Gounaris I, Ros S, Jimenez-Linan M, Garcia-Corbacho J, Patel K, Østrup O, Murphy S, Eldridge MD, Gale D, Stewart GD, Burge J, Cooper WN, van der Heijden MS, Massie CE, Watts C, Corrie P, Pacey S, Brindle KM, Baird RD, Mau-Sørensen M, Parkinson CA, Smith CG, Brenton JD, Rosenfeld N: Enhanced detection of circulating tumor DNA by fragment size analysis. *Sci. Transl. Med.* 10: eaat4921, 2018
170. Jovelet C, Ileana E, Le Deley M-C, Motte N, Rosellini S, Romero A, Lefebvre C, Pedrero M, Pata-Merci N, Droin N, Deloger M, Massard C, Hollebecque A, Ferte C, Boichard A, Postel-Vinay S, Ngo-Camus M, De Baere T, Vielh P, Scoazec J-Y, Vassal G, Eggermont A, Andre F, Soria J-C, Lacroix L: Circulating Cell-Free Tumor DNA Analysis of 50 Genes by Next-Generation Sequencing in the

Prospective MOSCATO Trial. *Clin. Cancer Res.* 22: 2960–2968, 2016

171. Guinney J, Dienstmann R, Wang X, de Reyniès A, Schlicker A, Soneson C, Marisa L, Roepman P, Nyamundanda G, Angelino P, Bot BM, Morris JS, Simon IM, Gerster S, Fessler E, De Sousa E Melo F, Missiaglia E, Ramay H, Barras D, Homicsko K, Maru D, Manyam GC, Broom B, Boige V, Perez-Villamil B, Laderas T, Salazar R, Gray JW, Hanahan D, Tabernero J, Bernards R, Friend SH, Laurent-Puig P, Medema JP, Sadanandam A, Wessels L, Delorenzi M, Kopetz S, Vermeulen L, Tejpar S: The consensus molecular subtypes of colorectal cancer. *Nat. Med.* 21: 1350–1356, 2015
172. Pardoll DM: The blockade of immune checkpoints in cancer immunotherapy. *Nat. Rev. Cancer* 12: 252–264, 2012
173. Smyrk TC, Watson P, Kaul K, Lynch HT: Tumor-infiltrating lymphocytes are a marker for microsatellite instability in colorectal carcinoma. *Cancer* 91: 2417–2422, 2001
174. Mohan S, Heitzer E, Ulz P, Lafer I, Lax S, Auer M, Pichler M, Gerger A, Eisner F, Hoefler G, Bauernhofer T, Geigl JB, Speicher MR: Changes in Colorectal Carcinoma Genomes under Anti-EGFR Therapy Identified by Whole-Genome Plasma DNA Sequencing. *PLoS Genet.* 10: e1004271, 2014
175. Strickler JH, Loree JM, Ahronian LG, Parikh AR, Niedzwiecki D, Pereira AAL, McKinney M, Korn WM, Atreya CE, Banks KC, Nagy RJ, Meric-Bernstam F, Lanman RB, Talasaz A, Tsigelny IF, Corcoran RB, Kopetz S: Genomic Landscape of Cell-Free DNA in Patients with Colorectal Cancer. *Cancer Discov.* 8: 164–173, 2018
176. Zill OA, Banks KC, Fairclough SR, Mortimer SA, Vowles J V, Mokhtari R, Gandara DR, Mack PC, Odegaard JI, Nagy RJ, Baca AM, Eltoukhy H, Chudova DI, Lanman RB, Talasaz A: The Landscape of Actionable Genomic Alterations in Cell-Free Circulating Tumor DNA from 21,807 Advanced Cancer Patients. *Clin. Cancer Res.* 24: 3528–3538, 2018
177. Berger AW, Schwerdel D, Welz H, Marienfeld R, Schmidt SA, Kleger A, Ettrich TJ, Seufferlein T: Treatment monitoring in metastatic colorectal cancer patients by quantification and KRAS genotyping of circulating cell-free DNA. *PLoS One* 12: e0174308, 2017
178. De Vlaminck I, Martin L, Kertesz M, Patel K, Kowarsky M, Strehl C, Cohen G, Luikart H, Neff NF, Okamoto J, Nicolls MR, Cornfield D, Weill D, Valantine H,

Khush KK, Quake SR: Noninvasive monitoring of infection and rejection after lung transplantation. *Proc. Natl. Acad. Sci.* 112: 13336–13341, 2015

179. Yu D, An G, Xu L: Investigation of Efficacy Evaluation Comparison of cfDNA and CEA in Colorectal Cancer. *Clin. Lab.* 62: 1947–1953, 2016
180. Osumi H, Shinozaki E, Takeda Y, Wakatsuki T, Ichimura T, Saiura A, Yamaguchi K, Takahashi S, Noda T, Zembutsu H: Clinical relevance of circulating tumor DNA assessed through deep sequencing in patients with metastatic colorectal cancer. *Cancer Med.* 8: 408–417, 2019
181. Overman MJ, Modak J, Kopetz S, Murthy R, Yao JC, Hicks ME, Abbruzzese JL, Tam AL: Use of Research Biopsies in Clinical Trials: Are Risks and Benefits Adequately Discussed? *J. Clin. Oncol.* 31: 17–22, 2013
182. VanderLaan PA, Yamaguchi N, Folch E, Boucher DH, Kent MS, Gangadharan SP, Majid A, Goldstein MA, Huberman MS, Kocher ON, Costa DB: Success and failure rates of tumor genotyping techniques in routine pathological samples with non-small-cell lung cancer. *Lung Cancer* 84: 39–44, 2014
183. Gerlinger M, Rowan AJ, Horswell S, Larkin J, Endesfelder D, Gronroos E, Martinez P, Matthews N, Stewart A, Tarpey P, Varela I, Phillimore B, Begum S, McDonald NQ, Butler A, Jones D, Raine K, Latimer C, Santos CR, Nohadani M, Eklund AC, Spencer-Dene B, Clark G, Pickering L, Stamp G, Gore M, Szallasi Z, Downward J, Futreal PA, Swanton C: Intratumor Heterogeneity and Branched Evolution Revealed by Multiregion Sequencing. *N. Engl. J. Med.* 366: 883–892, 2012
184. Russo M, Siravegna G, Blaszkowsky LS, Corti G, Crisafulli G, Ahronian LG, Mussolin B, Kwak EL, Buscarino M, Lazzari L, Valtorta E, Truini M, Jessop NA, Robinson HE, Hong TS, Mino-Kenudson M, Di Nicolantonio F, Thabet A, Sartore-Bianchi A, Siena S, Iafrate AJ, Bardelli A, Corcoran RB: Tumor Heterogeneity and Lesion-Specific Response to Targeted Therapy in Colorectal Cancer. *Cancer Discov.* 6: 147–153, 2016
185. Lebofsky R, Decraene C, Bernard V, Kamal M, Blin A, Leroy Q, Rio Frio T, Pierron G, Callens C, Bieche I, Saliou A, Madic J, Rouleau E, Bidard F-C, Lantz O, Stern M-H, Le Tourneau C, Pierga J-Y: Circulating tumor DNA as a non-invasive substitute to metastasis biopsy for tumor genotyping and personalized medicine in a prospective trial across all tumor types. *Mol. Oncol.* 9: 783–790, 2015

186. Kidess E, Heirich K, Wiggin M, Vysotskaia V, Visser BC, Marziali A, Wiedenmann B, Norton JA, Lee M, Jeffrey SS, Poultsides GA: Mutation profiling of tumor DNA from plasma and tumor tissue of colorectal cancer patients with a novel, high-sensitivity multiplexed mutation detection platform. *Oncotarget* 6: 2549–2561, 2015
187. Beije N, Helmijr JC, Weerts MJA, Beaufort CM, Wiggin M, Marziali A, Verhoef C, Sleijfer S, Jansen MPHM, Martens JWM: Somatic mutation detection using various targeted detection assays in paired samples of circulating tumor DNA, primary tumor and metastases from patients undergoing resection of colorectal liver metastases. *Mol. Oncol.* 10: 1575–1584, 2016
188. Kaisaki PJ, Cutts A, Popitsch N, Camps C, Pentony MM, Wilson G, Page S, Kaur K, Vavoulis D, Henderson S, Gupta A, Middleton MR, Karydis I, Talbot DC, Schuh A, Taylor JC: Targeted Next-Generation Sequencing of Plasma DNA from Cancer Patients: Factors Influencing Consistency with Tumour DNA and Prospective Investigation of Its Utility for Diagnosis. *PLoS One* 11: e0162809, 2016
189. McDonald BR, Contente-Cuomo T, Sammut S-J, Odenheimer-Bergman A, Ernst B, Perdigones N, Chin S-F, Farooq M, Mejia R, Cronin PA, Anderson KS, Kosiorek HE, Northfelt DW, McCullough AE, Patel BK, Weitzel JN, Slavin TP, Caldas C, Pockaj BA, Murtaza M: Personalized circulating tumor DNA analysis to detect residual disease after neoadjuvant therapy in breast cancer. *Sci. Transl. Med.* 11: eaax7392, 2019
190. Khan KH, Cunningham D, Werner B, Vlachogiannis G, Spiteri I, Heide T, Mateos JF, Vatsiou A, Lampis A, Damavandi MD, Lote H, Huntingford IS, Hedayat S, Chau I, Tunariu N, Mentrasti G, Trevisani F, Rao S, Anandappa G, Watkins D, Starling N, Thomas J, Peckitt C, Khan N, Rugge M, Begum R, Hezelova B, Bryant A, Jones T, Proszek P, Fassan M, Hahne JC, Hubank M, Braconi C, Sottoriva A, Valeri N: Longitudinal Liquid Biopsy and Mathematical Modeling of Clonal Evolution Forecast Time to Treatment Failure in the PROSPECT-C Phase II Colorectal Cancer Clinical Trial. *Cancer Discov.* 8: 1270–1285, 2018
191. Liu Z, Zhang Y, Niu Y, Li K, Liu X, Chen H, Gao C: A Systematic Review and Meta-Analysis of Diagnostic and Prognostic Serum Biomarkers of Colorectal Cancer. *PLoS One* 9: e103910, 2014
192. Ito K, Hibi K, Ando H, Hidemura K, Yamazaki T, Akiyama S, Nakao A: Usefulness

of analytical CEA doubling time and half-life time for overlooked synchronous metastases in colorectal carcinoma. *Jpn. J. Clin. Oncol.* 32: 54–8, 2002

193. Zhou J, Chang L, Guan Y, Yang L, Xia X, Cui L, Yi X, Lin G: Application of Circulating Tumor DNA as a Non-Invasive Tool for Monitoring the Progression of Colorectal Cancer. *PLoS One* 11: e0159708, 2016
194. Korphaisarn K, Morris VK, Overman MJ, Fogelman DR, Kee BK, Raghav KPS, Manuel S, Shureiqi I, Wolff RA, Eng C, Menter D, Hamilton SR, Kopetz S, Dasari A: *FBXW7* missense mutation: a novel negative prognostic factor in metastatic colorectal adenocarcinoma. *Oncotarget* 8: 39268–39279, 2017

7 APPENDIX

7.1 Publications derived from this work

Herrmann S, Zhan T, Betge J, Rauscher B, Belle S, Gutting T, Schulte N, Jesenofsky R, Härtel N, Gaiser T, Hofheinz R, Ebert MP, Boutros M: Detection of mutational patterns in cell-free DNA of colorectal cancer by custom amplicon sequencing. *Mol. Oncol.* 1878–0261.12539, 2019

7.2 Presentations related to this work

Herrmann S, Zhan T, Betge J, Belle S, Jesenofsky R, Rauscher B, Boutros M, Ebert M: Sequenzierung von Hotspotmutationen in zellfreier DNA (cirDNA) von Patienten mit gastrointestinalen Tumoren mittels modularen Amplicon Panels. *Z. Gastroenterol.* 54: KV375, 2016. Viszeralmedizin 2016, 71. Jahrestagung der Deutschen Gesellschaft für Gastroenterologie, Verdauungs- und Stoffwechselkrankheiten mit Sektion Endoskopie, 23.09.2016, Hamburg (short presentation).

Zhan T, Herrmann S, Betge J, Rauscher B, Belle S, Ebert M, Boutros M: Detection of mutations in circulating cell-free DNA using custom amplicon panels and deep sequencing. Personalized Medicine - International Symposium 2017, 03.07.2017, Heidelberg (presentation).

Herrmann S, Zhan T, Betge J, Rauscher B, Belle S, Boutros M, Ebert M: Detection of mutational patterns in circulating cell-free DNA of colorectal cancer by custom amplicon sequencing. Liquid Biopsies - EMBL Course 2019, 23.09.2019, Heidelberg (presentation).

7.3 Awards related to this work

MD scholarship for „Molekulare Gastroenterologische Onkologie“, Medical Faculty of Mannheim, Heidelberg University (2015-2016)

7.4 List of figures

7.5 List of tables

8 CURRICULUM VITAE

PERSONAL DATA

Name:	Herrmann, Simon
Date of birth:	26.02.1992
Place of birth:	Tübingen

SCHOOL

2002 – 2011	Eugen-Bolz-Gymnasium, Rottenburg am Neckar
10.06.2011	Internationales Abitur

UNIVERSITY

WS 2011/12	Start of Medical School at the University of Heidelberg
02.09.2013	1st state examination, „sehr gut“
12.10.2017	2nd state examination, “sehr gut”
21.12.2018	3rd state examination, „gut“ and graduation at the Medical Faculty Mannheim, Heidelberg University with the grade „sehr gut“.
26.12.2018	Approbation
2015-2019	Doctoral thesis at the Department of Medicine II, University Hospital Mannheim, Medical Faculty Mannheim, Heidelberg University

WORK EXPERIENCE

since 01.06.2019	Resident at the Department of Medicine II, University Hospital Mannheim, Heidelberg University

9 ACKNOWLEDGEMENTS

I thank Prof. Dr. Matthias Ebert and Prof. Dr. Michael Boutros for giving me the opportunity to work on this project. I thank Prof. Dr. Matthias Ebert for excellent supervision of this thesis and great guidance for my research projects.

I thank my two colleagues and project supervisors Dr. Tianzuo Zhan and Dr. Johannes Betge for their huge effort and support as well as the innumerable helpful discussions. I enjoyed our work as a team.

I thank Benedikt Rauscher for his work and endurance in bioinformatic issues, especially concerning the analysis of sequencing data. I thank Sabine Öttinger and Thilo Miersch for technical help with sequencing experiments. I thank Frank Herweck for his technical support in the lab and all other members of the M. Boutros lab and M. Ebert lab.
I thank Katharina Abkai for her work on the MALIBU biobank and all colleagues of the Department of Internal Medicine II for their effort to collect patient material.

I thank Prof. Dr. Timo Gaiser for providing samples and his help determining tumor cell contents of tissue samples. I thank Prof. Dr. Ralf Hofheinz and his colleagues of the Tagestherapiezentrum for the chance to include their patients into this project and their effort to collect patient material.

Finally, I want to thank my family and friends, especially Johanna who made the past years the best years of my life. Without your support, I would not have been able to succeed.