Medizinische Fakultät

der

Universität Duisburg-Essen

Aus dem Institut für Pathologie

Parallel progression of primary colorectal tumors and metastases

I n a u g u r a l d i s s e r t a t i o n

zur

Erlangung des Doktorgrades der Medizin

durch die Medizinische Fakultät

der Universität Duisburg-Essen

vorgelegt von

Jan Bernhard Schmeller

aus Wertheim

2020

DOI: 10.17185/duepublico/74312
URN: urn:nbn:de:hbz:464-20210611-075950-3

Dekan:	Herr Univ.-Prof. Dr. med. J. Buer
1. Gutachter:	Herr Univ.-Prof. Dr. med. K. W. Schmid
2. Gutachter:	Herr Univ.-Prof. Dr. med. St. Kasper-Virchow

Tag der mündlichen Prüfung: 24. März 2021

Publikationen zu der vorliegenden Dissertation:

Erstautorschaften:

1. Schmeller, J. *et al.* Setting out the frame conditions for feasible use of FFPE derived RNA. *Pathol Res Pract* 215, 381-386, doi:10.1016/j.prp.2018.12.027 (2019).
2. Mairinger, F. D.; Schmeller, J *et al.* Immunohistochemically detectable metallothionein expression in malignant pleural mesotheliomas is strongly associated with early failure to platin-based chemotherapy. *Oncotarget* 9, 22254-22268, doi:10.18632/oncotarget.24962 (2018).

Co-Autorschaften:

3. Borchert, S. *et al.* Gene expression profiling of homologous recombination repair pathway indicates susceptibility for olaparib treatment in malignant pleural mesothelioma in vitro. *BMC Cancer* 19, 108, doi:10.1186/s12885-019-5314-0 (2019).
4. Wessolly, Michael et al. "Processing Escape Mechanisms Through Altered Proteasomal Cleavage of Epitopes Affect Immune Response in Pulmonary Neuroendocrine Tumors" *Technology in Cancer Research & Treatment* vol. 17 1533033818818418. 13 Dec. 2018, doi:10.1177/1533033818818418
5. Walter, R. F. H. et al. Inhibition of MDM2 via Nutlin-3A: A Potential Therapeutic Approach for Pleural Mesotheliomas with MDM2-Induced Inactivation of Wild-Type P53. Journal of oncology 2018, 1986982, doi:10.1155/2018/1986982 (2018).
6. Mairinger, F. D. et al. miRNA regulation is important for DNA damage repair and recognition in malignant pleural mesothelioma. Virchows Arch 470, 627-637, doi:10.1007/s00428-017-2133-z (2017).
7. Walter, R. F. *et al.* ACTB, CDKN1B, GAPDH, GRB2, RHOA and SDCBP Were Identified as Reference Genes in Neuroendocrine Lung Cancer via the nCounter Technology. *PLoS One* 11, e0165181, doi:10.1371/journal.pone.0165181 (2016).
8. Walter, R. F. et al. Massive parallel sequencing and digital gene expression analysis reveals potential mechanisms to overcome therapy resistance in pulmonary neuroendocrine tumors. J Cancer 7, 2165-2172, doi:10.7150/jca.16925 (2016).
9. Walter, R. F. et al. Screening of Pleural Mesotheliomas for DNA-damage Repair Players by Digital Gene Expression Analysis Can Enhance Clinical Management of Patients Receiving Platin-Based Chemotherapy. J Cancer 7, 1915-1925, doi:10.7150/jca.16390 (2016).

Index

1. Introduction

Colorectal cancer (CRC) is a biologically heterogeneous disease with infaust prognosis. In 2014, 25.000 patients died from CRC related causes in Germany. The incidence doubles every decade for individuals above 40 years of age (Herold, 2018). CRC is a malignant transformation of cells localized in the colon or the rectum. However, they are often subsumed together due to their similar features. CRC pathogenesis depends on tumor localization and is different between the proximal colon (right side) and distal colon (left side) (Figure 1). Furthermore, right sided and left sided colorectal cancer patients show distinct responses to therapy (Baran et al., 2018). CRC has a high mortality rate at late stages (Baran et al., 2018). Symptoms may include unusual bowel habits, such as continuous diarrhea or constipation, blood in the stool, rectal bleeding and abdominal pain.

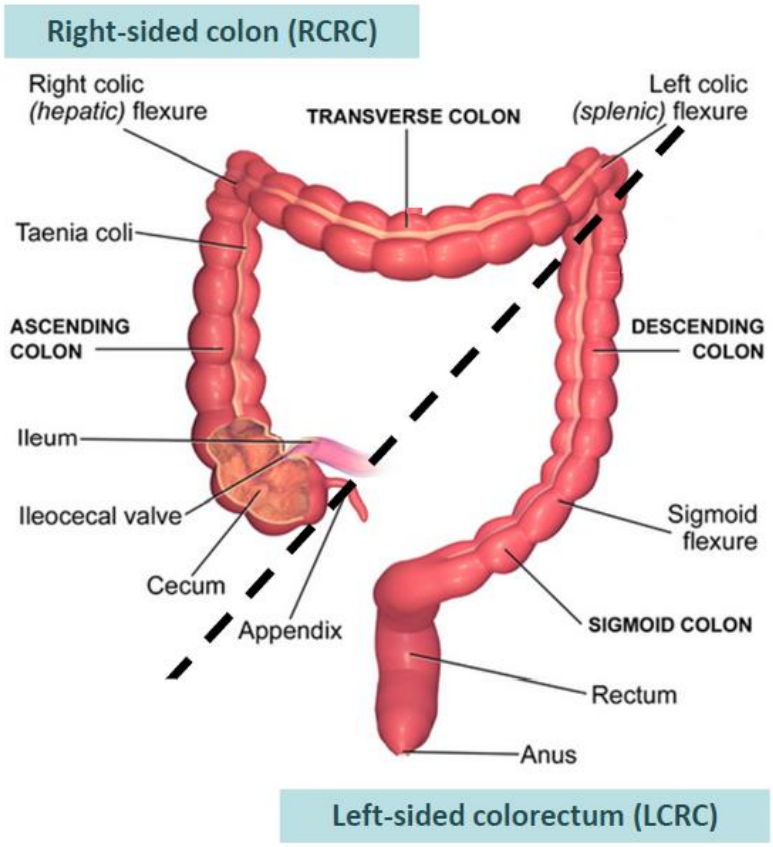

Figure 1: Illustration of the large Intestine. Front view of the abdomen (Blausen.com, 2014).

1.1 Epidemiology of CRC

CRC is the most common cancer type worldwide in both genders with an incidence rate of 10.2% and a mortality rate of 9.2% (F. Bray et al., 2018). CRC is the third deadliest cancer in the US, outranked by prostate and lung cancer in men and breast and lung cancer in women. 8% of all new cancer cases in the United States in 2016 were CRCs. The incidence rate of CRC in 2018 worldwide was higher in men than in women with a rate of 23.6 and 16.3 per 100.000 population, respectively (Rawla et al., 2019). The number of CRC cases decreased over the last decades and the 5-year survival rate is improving (Marley & Nan, 2016). CRC rates increase with age; people older than 50 years have increased CRC risk of up to 90%.

The overall survival (OS) is strongly influenced by the presence of metastasis and is related to the stage of cancer. The prognosis for right sided colorectal (RCRC) cancer is worse than for left sided colorectal cancer (LCRC) (Baran et al., 2018). CRC can be located in the descending colon (40% to 42%), in the rectosigmoid and rectum (30% to 33%), cecum and ascending colon (25% to 30%) and transverse colon (10% to 13%) (Ferri, 2019). The large intestine is depicted in Figure 1.

The most common risk factors are advanced age (>50), lack of exercise, obesity, increased meat and energy uptake, fat or alcohol consumption, low fiber diet, family history of CRC and smoking (Marley & Nan, 2016). Persons with rare inherited conditions like familial adenomatous polyposis, hereditary nonpolyposis colorectal cancer, attenuated familial adenomatous polyposis, juvenile polyposis syndrome and Peutz-Jeghers syndrome also have a higher risk of developing CRC. 1% of all CRC is related to chronic inflammation associated with ulcerative colitis (Yashiro, 2014).

Metastases in CRC patients are the major cause of cancer mortality, since metastatic CRC is to date mostly incurable and often already present at initial diagnosis (Vatandoust et al., 2015). Manifestation of metastases is observed in every second CRC patient. Resection of the metastasis can be curative; unresectable metastasis leads to poor prognosis. The prognosis for patients with metastatic CRC depends on several factors such as the metastatic site: oligometastatic disease which is limited in number and distribution and can be treated locally and diffuse metastatic disease which requires systemic treatment and is often

incurable (Niibe & Hayakawa, 2010). Differences between colon and rectal cancers regarding metastatic sites have been observed. Generic adenocarcinoma metastasizes more frequently into the liver and less within the peritoneum compared with mucinous and signet ring adenocarcinomas. Thoracic organs and the nervous system are the most common metastatic sites of rectal cancer; the peritoneum is a less common metastatic site. Peritoneal metastases frequently occur with ovarian and pleural metastases. Metastases in the lungs often appear together with nervous system metastases (Riihimaki et al., 2016). During disease progression, half of the CRC patients will develop liver metastasis and one third of the patients with metastatic CRC have liver as a sole metastatic site. The high frequency of liver metastases compared to metastases in other organs is believed to be due to the venous drainage of the colon and rectum (Vatandoust et al., 2015).

1.2 CRC pathogenesis

The acquisition of genetic and epigenetic alterations over time leads to CRC pathogenesis via the progression of normal colorectal cells to adenomas and the subsequent transformation of adenomas to carcinomas (Fearon & Vogelstein, 1990; WM, 2005). This sequence of events leads to most of CRC originated in the colon ascendance and is named colorectal adenoma-carcinoma (Figure 2). Typical gene alterations that are involved in this process are *APC* gene loss, *KRAS* mutation, *DCC* gene and *TP53* gene loss. APC is a tumor suppressor protein and an antagonist of the Wnt signaling pathway (Testa et al., 2018). *KRAS* proto-oncogene translates for K-ras protein, a GTPase which when mutated leads to various malignancies (Testa et al., 2018). The tumor suppressor protein DCC is a single transmembrane receptor (Duman-Scheel, 2012). *TP53* is a tumor suppressor gene which encodes for a protein involved in cell cycle, apoptosis, and genomic stability (Testa et al., 2018). Adenoma class I formed after *APC* gene loss is characterized by low-grade dysplasia. Adenoma class II formed from adenoma class I with *KRAS* mutation is also characterized by low-grade dysplasia whereas, adenoma class III formed after *DCC* gene loss in adenoma class II is characterized by high-grade dysplasia. The carcinoma results from *TP53* gene loss in adenoma class III and can either be low-grade (high or moderate differentiation) or high-grade (low or undifferentiated) (Herold, 2018).

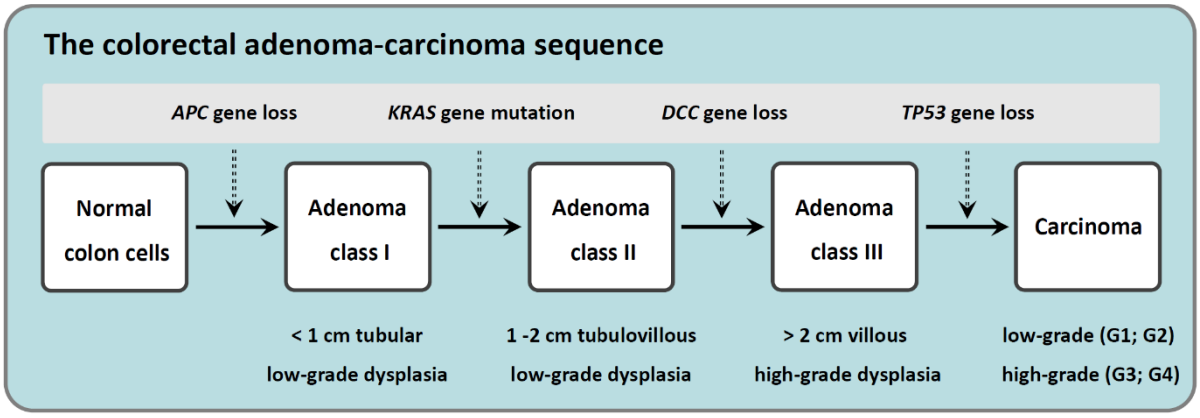

Figure 2: The colorectal adenoma-carcinoma sequence. Adenomatous polyposis coli (*APC*) gene is often deleted in 5q21 leading to adenoma class I, where mutation in KRAS Proto-Oncogene (*KRAS*) can lead to adenoma class II. In adenoma class II, 18q21 carrying deleted in colorectal carcinoma (*DCC*) gene is often deleted leading to adenoma class III, where *TP53* loss causes carcinoma. 2-5% of adenocarcinomas occur as multiple primary tumors. Low-grade carcinomas: G1 = high differentiation; G2 = moderate differentiation. High-grade carcinomas: G3 = low differentiation; mucinous and non-mucinous adenocarcinoma; G4 = Signet ring cell carcinoma, small cell and undifferentiated carcinoma. High-grade carcinomas show early lymphatic metastasis. (Wu et al., 2018). Adapted from Herold 2018 (Herold, 2018).

1.3 Diagnosis of CRC

Diagnosis of CRC either occurs from an assessment of symptomatic patients, or as a result of a routine screening. Typical symptoms of CRC include abdominal pain, blood in stools, rectal bleeding and changes in bowel habits. Furthermore, fatigue, shortness of breath, and unintended weight loss can also be related symptoms. Asymptomatic patients are diagnosed at a pre-clinical stage during a screening for CRC.

Common non-invasive diagnostic methods include fecal occult blood test, immunohistochemical fecal occult blood test and analyzing of tumor markers (e.g. Carcinoembryonic antigen (CEA), Carbohydrate-Antigen 19-9 (CA 19-9), Tumor-associated glycoprotein 72 (TAG-72). Colonoscopy is the diagnostic and screening standard and provides high accuracy. Via colonoscopy localization of the tumor and biopsy sampling are simultaneously achieved. Endoscopic techniques such as chromoendoscopy and magnification endoscopy, CT colonography and detection of relevant biomarkers are further available diagnostic tools (Kuipers et al., 2015). Additional diagnostic methods for

metastasis are ultrasound, magnetic resonance imaging, x-ray, positron emission tomography scan (Swiderska et al., 2014). Following colonoscopy, histological and molecular profiling of the tumor enabled by biopsy sampling determines the therapeutic strategy.

In World Health Organization (WHO) histologic classification of CRC, a number of carcinomas are listed such as mucinous, signet ring cell, medullary, micropapillary, serrated, cribriform comedo-type, adenosquamous, spindle cell, squamous cell and undifferentiated. The majority of colorectal carcinomas, namely more than 90% are adenocarcinomas originating from epithelial cells of the colorectal mucosa (Aaltonen et al., 2000). Histologic tumor grading is based on glandular formation. Well differentiated adenocarcinoma shows >95% gland formation, moderately differentiated adenocarcinoma 50-95% and poorly differentiated adenocarcinoma <50% gland formation. Most of colorectal adenocarcinomas (~70%) are diagnosed as moderately differentiated, while 10% as diagnosed as well differentiated and 20% as poorly differentiated (M. Fleming et al., 2012). Mucinous adenocarcinoma can be diagnosed when extracellular mucins are over 50% of tumor areas and signet ring cell carcinoma when signet ring cells are over 50% of tumor component. Medullary carcinoma is diagnosed when tumor cells appear as solid or sheet-like structures and lymphocytic infiltration is prominent with tumor-infiltrating lymphocytes and neutrophils. Micropapillary adenocarcinoma is another rare histologic type that is diagnosed by small tumor cell clusters which are surrounded by empty spaces. Lymphatic and vascular invasion are characteristics of micropapillary adenocarcinoma that are considered to be poor prognosis factors (B. H. Kim et al., 2020). When an area of definite squamous differentiation is present in the tumor then adenosquamous carcinoma is diagnosed (B. H. Kim et al., 2020). Serrated adenocarcinoma is characterized by neoplastic glands with prominent epithelial serrations, eosinophilic and abundant cytoplasm, low nucleus-to-cytoplasm ratio, and vesicular nuclei. It is morphologically similar to serrated polyp (Tuppurainen et al., 2005).Undifferentiated carcinoma is diagnosed When the epithelial tumor shows no molecular, morphological and immunohistochemical evidence of specific differentiation then it is diagnosed as undifferentiated (B. H. Kim et al., 2020).

In addition to histologic classification, histologic staging is also an essential part of diagnosis that enables prognosis prediction and choice of treatment. The staging system most often

used for colorectal cancer is the Union for International Cancer Control (UICC) TNM system. The "T" stands for tumor and is used to describe how deeply the primary tumor has grown into the lining of colon or rectum. The "N" stands for lymph nodes and describes the spread of the tumor to regional lymph nodes. The "M" describes metastasis to distant sites. The results of TNM classification are combined to determine the stage: There are 5 stages: stage 0 (zero) and stages I through IV with subcategories. Detailed description of TNM staging is shown below in Table 1.

Table 1: TNM classification of malignant tumors for CRC (Union Internationale Contre le Cancer (UICC)) (Brierley et al.).

TNM Classification (UICC. 2010):					
T_{is}	Carcinoma in situ	M_0	No distant metastasis		
T_1	Submucosa	M_{1a}	Distant metastasis in one organ		
T_2	Muscularis propria	M_{1b}	Distant metastasis in more than one organ		
T_3	Subserosa	Stage 0	Tis	N0	M0
T_{4a}	Visceral peritoneum	Stage I	T1,T2	N0	M0
T_{4b}	Adjacent organ or structure	Stage II	T3,T4	N0	M0
N_0	No regional metastatic lymph nodes (LN)	Stage III	Any T	N1,N2	M0
$N_{1a/b}$	1 LN/ 2 - 3 LN	Stage IV	Any T	Any N	M1
N_{1c}	Tumor in fatty tissue of the subserosa; no regional LN				
N_{2a}	4 - 6 LN				
N_{2b}	≥ 7 LN				

A further technique widely used in CRC diagnosis in immunohistochemistry. CRCs express nuclear transcription factor CDX2, which is highly specific for intestinal epithelial cells and colorectal cancers. Other markers used for identification of colorectal adenocarcinomas include: CEA, CA19-9, Calretinin, Keratin 20 (CK20) and Keratin 7 (CK7), as well as Mucin (MUC1/2) (Taliano et al., 2013). Additionally, mutations in the oncogenes KRAS, BRAF, and PIK3CA can also be used as markers for many colorectal cancers. An effective marker for well-differentiated cancers is a membranous protein which is also a target for antibody radioimmunotherapy and is encoded by *GPA33* (Baptistella et al., 2016).

1.4 Treatment of CRC

The most common treatment for primary disease is laparoscopic surgery for tumor resection where parts of the healthy surrounding tissue also get removed in order to ensure the total removal of the cancerous tissue. Resection of metastases affecting for example the liver and lungs is also common practice. Radiation therapy is frequently used, especially for rectal

cancer but also for some forms of metastatic disease. Neoadjuvant therapy (administration of therapeutics before the surgical treatment) and palliative chemotherapy are also available. (Heemskerk-Gerritsen et al., 2015; Hurwitz et al., 2004; Papamichael et al., 2015). In order to reduce the rate of local recurrence, neoadjuvant therapy with radiation before surgery or with a combination of radiation and chemotherapy are recommended for intermediate-stage and advanced-stage rectal cancer. For colon cancer there is no widely accepted neoadjuvant therapy (Kuipers et al., 2015). In contrast to neoadjuvant treatment, adjuvant therapy is used after primary treatments, such as surgery, to reduce the possibility of recurrence. Possible treatments used as adjuvant therapies are chemotherapy, radiation, immunotherapy and targeted therapy (Carrato, 2008).

Typical chemotherapeutics used in the treatment of CRC are 5-fluorouracil with leucovorin in combination with irinotecan (FOLFIRI); oxaliplatin (FOLFOX), capecitabine and oxaliplatin (XELOX) or 5-fluorouracil, leucovorin, irinotecan and oxaliplatin (FOLFOXIRI) (Porru et al., 2018). Targeted therapy is the foundation of precision medicine and an important tool for cancer treatment where molecular features of tumors are targeted in order to achieve the best possible outcome. Most targeted therapies consist of small-molecule drugs or monoclonal antibodies. For example, the monoclonal antibody bevacizumab inhibits the production of new blood vessels via inhibition of the vascular endothelial growth factor A (VEGFA) (Garcia-Romero et al., 2020). Cetuximab inhibits proliferation of cancer cells via epidermal growth factor receptor (EGFR) blockade (Moiseyenko et al., 2018). In addition, for microsatellite instability high (MSI-H) or mismatch repair deficient (dMMR) CRC, immune checkpoint inhibitors such as pembrolizumab and nivolumab are used off-label; they target programmed death-ligand 1 (PD-L1) and programmed cell death protein 1 **(**PD-1) respectively. Thereby, they stimulate the hosts immune system against the tumor (Huyghe et al., 2020).

Despite advances in CRC therapy, mortality rates have not substantially improved in the past decades. Further individualization of treatment and further development of screening strategies should contribute to increased survival rate.

1.5 The models of metastasis

Linear progression model

The basic characteristics of a linear model of metastatic seeds are the gradual progression of morphological malformations and the gradual accumulation of genetic alterations in the primary. Cancer cells of the primary tumor start proliferating after undergoing multiple rounds of mutation and therefore selection. Thereby, cancer cell clones migrate to other sides to initiate the growth of secondary tumors (Tan et al., 2015). Primary tumor size has been associated to the clonal expansion of fully metastatic clones. For instance, mutations of tumor suppressor gene *TP53* in breast cancer occur more frequently when the tumor reaches a size beyond 2 cm (Rivlin et al., 2011). Surgical intervention prior to this e.g. in smaller tumors would remove the neoplasia before this genetic event, and thereby prevent the migration of malignant cells to other sites. However, not only tumor cells diffusing at a late stage of primary tumor progression have the possibility of causing metastasis, but also tumor cells that diffuse earlier and are highly similar to the primary tumor cells. This model also describes the emergence of metastasis from metastasis (Fearon & Vogelstein, 1990; Klein, 2009).

Parallel progression model

In contrast to the linear progression model, where metastases occur after acquisition of the majority of mutations of the primary tumor cells, the parallel progression model describes an accumulation of mutations in a parallel manner and independently within the primary tumor and the metastasis due to tumor cell migration (Klein, 2009). In this model, a late clinical onset of symptoms is observed, resulting to late diagnosis of the primary tumor. Both models are depicted in Figure 3.

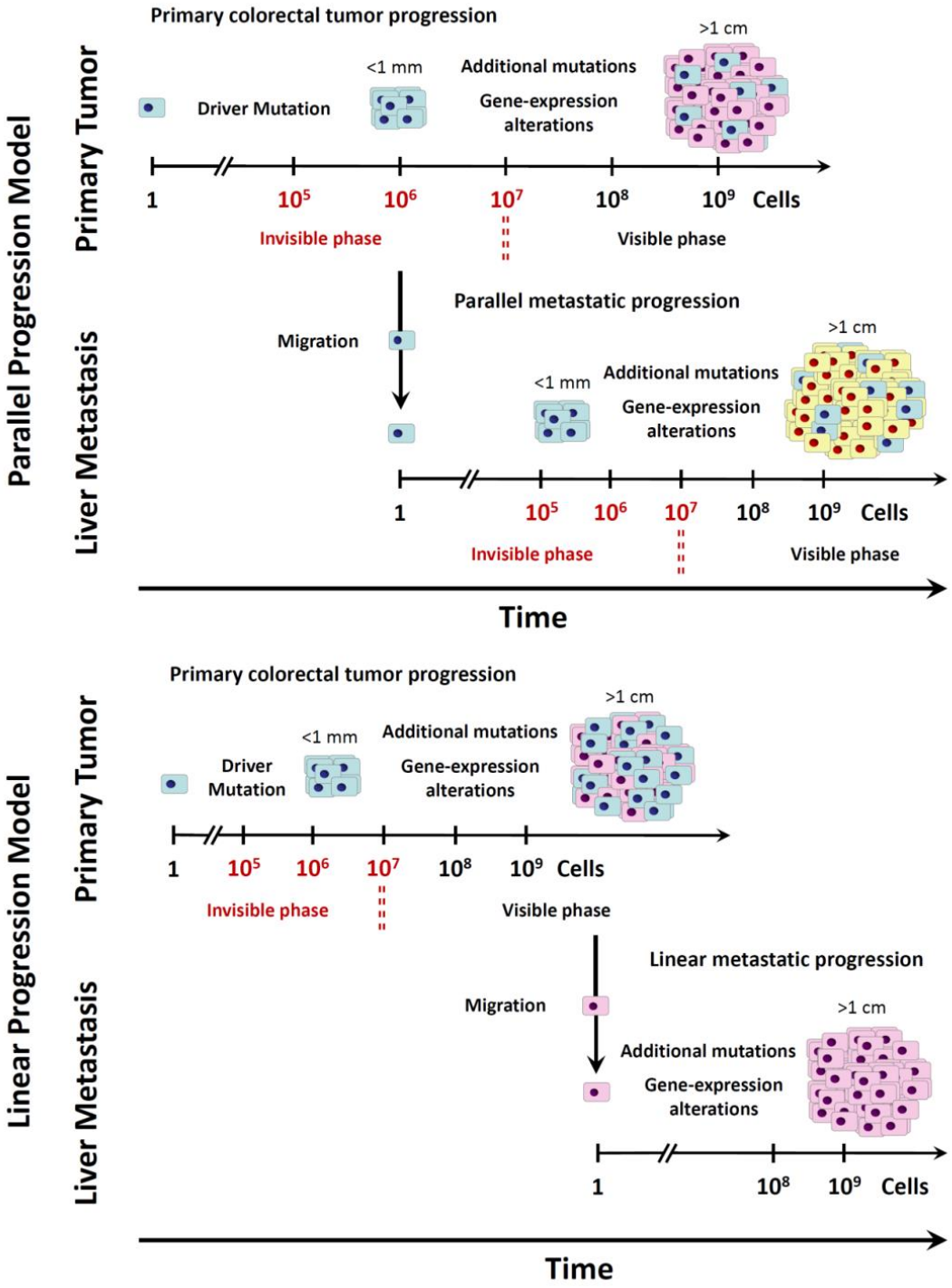

Figure 3: Comparison between the parallel progression model and the linear progression model (colorectal cancer). Adapted from Caswell (Caswell & Swanton, 2017) and Klein (Klein, 2009).

1.6 Genetic alterations in CRC

CRC is a very heterogeneous disease with a high variety of genetic and epigenetic alterations (Galon et al., 2014). Two major pathways for carcinogenesis have been described: the microsatellite instability (MSI) pathway and the chromosome instable (CIN) pathway. The MSI results in hypermutated tumors, mostly presenting in right sided CRC (Yaeger et al., 2018) (Baran et al., 2018). On the other hand, in the CIN pathway, alterated, non-hypermutated tumors are more common in left sided CRC (Baran et al., 2018). CpG island methylator phenotype (CIMP), is a subset of colorectal cancers based on the abnormal methylation of several genes and is occasionally observed in DNA mismatch repair pathway deficient as well as MSI-high tumors. However, CIMP definitions are yet to be clarified in order to elucidate whether CIMP is a universal phenomenon across various types of tumors (Hughes et al., 2013; Jia et al., 2016).

Microsatellite instability (MSI)

In the MSI pathway, a deficient DNA mismatch repair (MMR) system develops due to mutations or transcriptional silencing of *MMR* genes such as *MLH1*, *MSH2, MSH6,* and *PMS2* (Arnold et al., 2005). Identifying MMR deficiencies is of great importance for CRC therapy management since MSI tumors are phenotypically distinct and have different prognosis compared to other tumors (Sinicrope & Sargent, 2012). The most common gene presenting mutations (point mutations and insertions) is the tumor suppressor gene *MLH1*. *BRAF*, an oncogene, is involved in cell proliferation and is activated via a point mutation at codon 600 in 46% of the cases. Tumor suppressor genes like *MSH2, MSH3* and *MSH6* are mutated (point mutations, insertions) and thus inactive in 40% of the cases. In addition, *ACVR2A,* a further tumor suppressor gene, is involved in cellular growth and is deficient in 62% of the cases. *APC and TGFβR2* are tumor suppressor genes involved in Wnt signaling that have been detected to be altered in 51% of the cases (Baran et al., 2018). A table listing the function and the mutation rate of the aforementioned genes is shown below (Table 2).

Table 2: List of genes mutated in microsatellite instability (MSI) -high tumors

MSI-high tumors			
Gene		**Function**	**Mut. rate [%]**
ACVR2A	Activin receptor type-2A	Cellular growth	62
APC	Adenomatous-polyposis-coli-Protein	Inhibition of Wnt signaling	51
BRAF	Proto-oncogene B -rapidly accelerated fibrosarcoma	Proliferation and Survival	46
MLH1	MutL homolog 1	DNA mismatch repair	77
MSH 3	MutS Homolog 3	DNA mismatch repair	40
MSH2	DNA mismatch repair protein Msh2	DNA mismatch repair	40
MSH6	mutS homolog 6	DNA mismatch repair	40
POLE	DNA polymerase epsilon catalytic subunit	DNA proofreading	10
TGFβR2	Transforming growth factor, beta receptor II	Regulation of Wnt signaling	51

Chromosome instability

The CIN pathway describes a stepwise carcinogenesis initiated with the inactivation of *APC* tumor suppressor gene, which is then followed by the activation of *KRAS* oncogene, inactivation of *TP53* tumor suppressor gene and further mutations. According to the Cancer Genome Atlas study these are the most frequently mutated genes in the CIN pathway: *APC, TP53, KRAS, PIK3CA, FBXW7, SMAD4, TCF7L2,* and *NRAS.* The *APC* gene of which the function was described above is the most commonly mutated gene in CIN tumors (70-80% of the cases), followed by *TP53* with a mutation frequency of 50-60%. *KRAS and PIK3CA* are oncogenes which play a role in cell proliferation and survival. *KRAS* is found to be mutated in 40% of the tumors with point mutations in codons 12, 13 and 61. *PIK3CA* is mutated in 15-25% of the cases with point mutations in exon 9 and 20. *FBXW7,* involved in targeting oncoproteins for degradation is altered in 20% of the cases. *SMAD4,* a tumor suppressor gene which functions as an intracellular signal transmitter of TGF-β pathway is mutated in 10-15% of the tumors. *SMAD2* belongs to the same family as *SMAD4* but is found mutated only in a frequency of 5-10%. *NRAS* which plays a role in cell proliferation and cell survival is mutated by point mutations in codons 12, 13 and 61 in a frequency <5%. Further genes such as *CTNNB1, SOX9* and *ARIDIA* are also found mutated in CIN tumors but to a lower extent. The CIN phenotype is also observed when the expression of cell division regulator Aurora kinases and POLO-like kinases (Plks) is dysregulated (Baran et al., 2018; Cancer Genome Atlas, 2012). A simplified overview of the CRC pathway is depicted in Figure 4. A table listing the function and the mutation rate of the aforementioned genes is shown below (Table 3).

Table 3: List of genes mutated in chromosomal instability (CIN) tumors

CIN tumors			
Gene		**Function**	**Mut. rate [%]**
APC	Adenomatous-polyposis-coli-Protein	Inhibition of Wnt signaling	70-80
ARID1A	AT-rich interactive domain-containing protein 1A	Chromatin remodeling	5
ATM	ATM serine/threonine kinase	Cell cycle arrest	7
CTNNB1	Catenin beta 1	Tumor growth and invasion	< 5
FAM123B	Family With Sequence Similarity 123B	Causing mesenchymal phenotype	7
FBXW7	F-box/WD repeat-containing protein 7	Targets oncoproteins for degradation	20
KRAS	KRAS Proto-Oncogene, GTPase	Proliferation and survival	40
NRAS	NRAS proto-oncogene, GTPase	Proliferation and surival	<5
PIK3CA	PI3-Kinase P110 Subunit Alpha	Proliferation and survival	15-25
SMAD2	SMAD Family Member 2	Signal transmitter of TGF-β pathway	5-10
SMAD4	SMAD Family Member 4	Signal transmitter of TGF-β pathway	10-15
SOX9	SRY-Box Transcription Factor 9	Proliferation	4
TCF7L2	Transcription factor 7-like 2	Regulation of Wnt signaling	5
TP53	Tumor protein p53	Cell cycle arrest, apoptosis	50-60

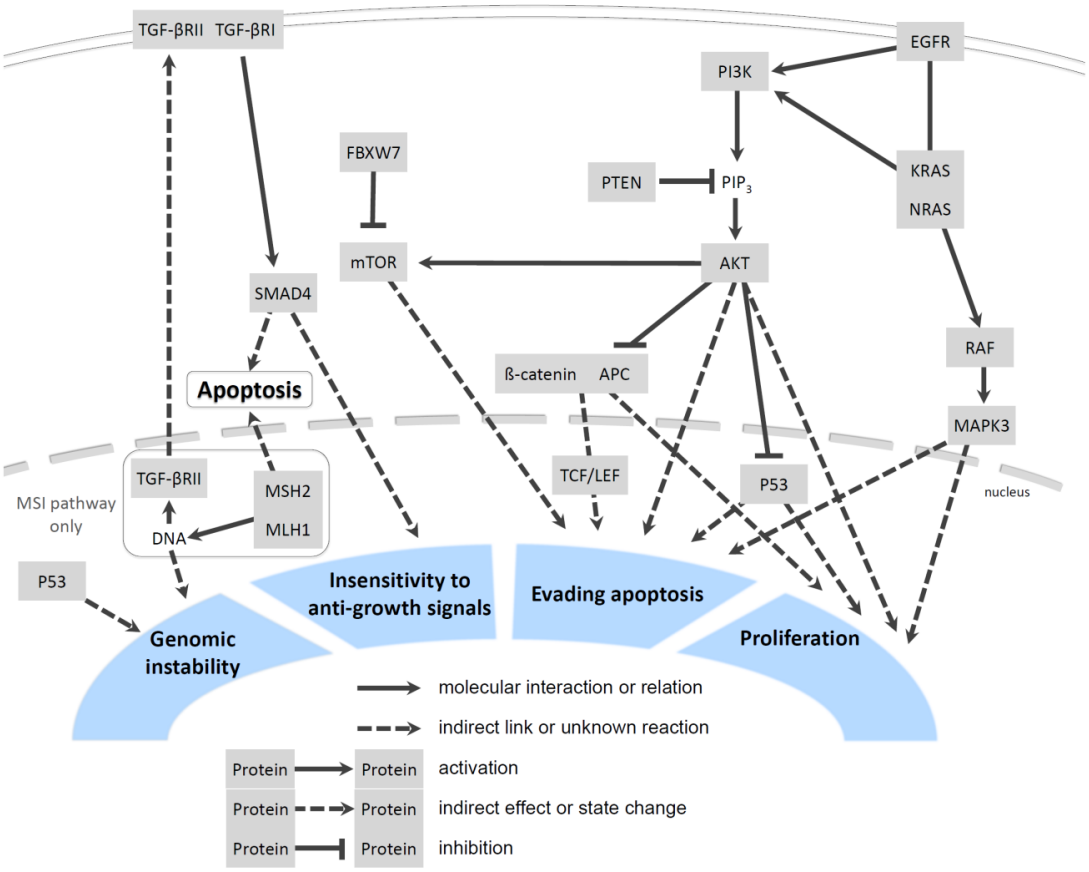

Figure 4: Simplified CRC pathway of CRC Chromosome unstable (CIN) pathway microsatellite unstable pathway. Adapted from Kyoto Encyclopedia of Genes and Genomes (KEGG): 05210 Colorectal cancer - Homo sapiens (human).

1.7 Effect of the metastasis models in therapy and clinical outcome

Elucidating the mechanisms of heterogeneity within tumors but also between tumors in primary and metastatic sites has led to increased understanding of therapy resistance development and relapse, and improved therapeutic outcomes (Caswell & Swanton, 2017).

Linear progression is a rather simple model for developing therapeutic strategies. At diagnosis, the disseminating tumor cells correspond to the primary tumor, representing the most aggressive clone.

Therapy selection for tumors within the parallel progression model is quite complicated compared to the linear progression tumors. Namely, the genomes of the primary tumor and the metastasis sites may vary significantly between each other, therefore, the choice of a targeting strategy can remain challenging (Klein, 2009; Koscielny & Tubiana, 2010).

2. Aims

For describing CRC progression from normal colon epithelia via likely benign adenoma to fully malignant carcinoma there is a defined set of genetic and histologic markers. These markers play an important role in the development of individual therapeutic strategies since they are of great significance for diagnosis, tumor progression, metastasis and biological phenotype (Baran et al., 2018; Marley & Nan, 2016).

However, the biological mechanisms of the development of either synchronous or metachronous metastasis are not yet fully understood.

To overcome this unsatisfactory situation, the present study was designed to:

1. discover mutation patterns as well as novel mutations in primary tumors and metastases in order to increase the understanding of molecular processes involved in CRC carcinogenesis, metastasis and tumor heterogeneity.
2. classify cases according to the progression of metastases in linear or parallel model.
3. identify signaling molecules and pathways involved in the hallmarks of cancer in order to be able to predict a potential metachronous metastasis development. This could lead to an adaption of clinical management of patients showing a high risk of time shifted metastases development due to changes in therapeutic strategies.
4. define, whether tumors and metastases should be treated differently by investigating tumor heterogeneity, which is of great relevance regarding therapeutic strategies and clinical outcome.
5. validate all findings with an independent cohort acquired from TCGA database.

Altogether, this study was designed to improve the current knowledge about colorectal cancer at a molecular level and to influence the choice of a therapeutic strategy by clinicians.

3. Material and Methods

3.1 Ethics approval

The use of material and retrospective data collection was performed with the approval of the local ethics committee of the Semmelweis University (TUKEB 83/2009) and included a waiver for informed consent. The study was approved by the ethics committee of the Medical Faculty University Hospital Essen, University Duisburg-Essen, Germany (Ethics approval 16-7075-BO). The investigation conforms to the principles outlined in the Declaration of Helsinki, and all patient samples were rendered anonymous.

3.2 Patient characteristics

3.2.1 Cohort A

Tumor samples from 30 patients with CRC and corresponding liver metastasis from two different institutions (Semmelweis University, University Hospital Essen) were analyzed (Cohort A). Primary tumor was resected between the years 2006-2012. Liver metastases were resected between the years 2006 and 2018. The patients were diagnosed at a mean age of 63 years.

Retrospectively collected formalin-fixed, paraffin-embedded (FFPE) tissues were used for the present study. The degree of histological differentiation, lymphatic invasion, localization and the quantity of necrosis was determined. A detailed overview of the clinicopathological factors is depicted in Table 4 and Table 5. Data on patient age, gender, primary tumor site, corresponding liver metastatic site, pathological stage and chemotherapy were retrieved from patient medical records. The clinicopathological factors were assessed according to the tumor node metastasis (TNM) classification of the Union internationale contre le cancer (UICC) classification of malignant tumors (Brierley et al.). A metastasis was defined as synchronous if the metastasis was operated within one year or the initial TNM was M1. A description of the TNM classification is depicted in Table 1. A metastasis was defined as metachronous if it occurred one year after primary resection.

Table 4: Clinicopathological characteristics of the patient cohort A

Total number of patients				30	100%
Age at diagnosis (years; mean±standard deviation)				63±11.7	
Gender	Male			24	80%
	Female			6	20%
Localization	Primary colorectal cancer	Rectum		11	37%
		Sigmoid colon		9	30%
		Transverse colon		2	7%
		Splenic flexure		1	3%
		Hepatic flexure		1	3%
		Cecum		1	3%
		Ascending colon		5	17%
	Corresponding liver metastasis	Left liver (Segment I; II; III; IV)		18	60%
		Right liver (Segment V; VI; VII; VIII)		7	23%
		Affecting both sides		3	10%
		(NOS[1])		2	7%
Stage at diagnosis		I		3	10%
		II		6	20%
		III		9	30%
		IV		12	40%
TNM at diagnosis[2]	T-stage	T2		3	10%
		T3		25	83%
		T4		2	7%
	N-stage	N0		11	37%
		N1		9	30%
		N2		9	30%
		NX		1	3%
Treatment	Neoaduvant chemotherapy		Yes	5	17%
			No	19	63%
	Adjuvant chemotherapy		Yes	21	70%
			No	1	3%
	No data available			8	27%
Metastasis	Synchronous liver metastases			17	57%
	Metachronous liver metastases			13	43%
	Number of metastasic sites		1	17	57%
			2	5	17%
			3	1	3%
			Multiple	7	23%

[1] Not Otherwise Specified (NOS)
[2] Tumor Node Metastasis (TNM)

Table 5: Histopathological characteristics of the patients in cohort A

Histopathological characteristics (Cohort A)				
Histotype	Primary Tumor	Adenocarcinoma (NOS[1])	24	80%
		Mucinous differentiation	2	7%
		(20%) mucinous (80%) adenocarinoma	1	3%
		N/A	3	10%
	Metastasis	Adenocarcinoma (NOS[1])	25	83%
		Mucinous differentiation	4	13%
		Signet ring cell differentiation	1	3%
Necrosis	Metastasis	-	4	13%
		+	14	47%
		++	7	23%
		+++	5	17%

[1] Not Otherwise Specified (NOS)

3.2.2 Cohort B

A second cohort (cohort B N= 66) was designed using data from Brannon et al.(Brannon et al., 2014). The colorectal adenocarcinoma data and the corresponding liver metastasis data were retrieved from cBioPortal (MSKCC, Genome Biol 2014). Cases with additional metastasis like ovary or lung metastasis were excluded.

3.3 KRAS mutation status

KRAS mutations of the Semmelweis patients (N=22) were identified at the Semmelweis University by microcapillary-based restriction fragment length analysis as described (Cserepes et al., 2014). Briefly, tumor-rich microscopic area stained with hematoxylin and eosin (H&E) was determined by pathologists prior to macrodissection from FFPE tissue or cytological smears. DNA was extracted using the MasterPure™ DNA Purification Kit (Epicentre Biotechnologies, WI) according to the instructions of the manufacturer. The microfluid-based restriction fragment detection system was characterized by 5% mutant tumor cell content sensitivity. Density ratio of the mutated band to the wild-type (WT) was calculated and samples containing >5% of the non-WT band were considered mutation positive based on the sensitivity threshold. The base-pair substitution in the mutant samples was verified and determined by sequencing using the ABI 3130 Genetic Analyzer System (Life Technologies, Carlsbad USA) with the BigDye® Terminator v1.1 Kit.

3.4 Targeted Amplicon Sequencing – Cohort A

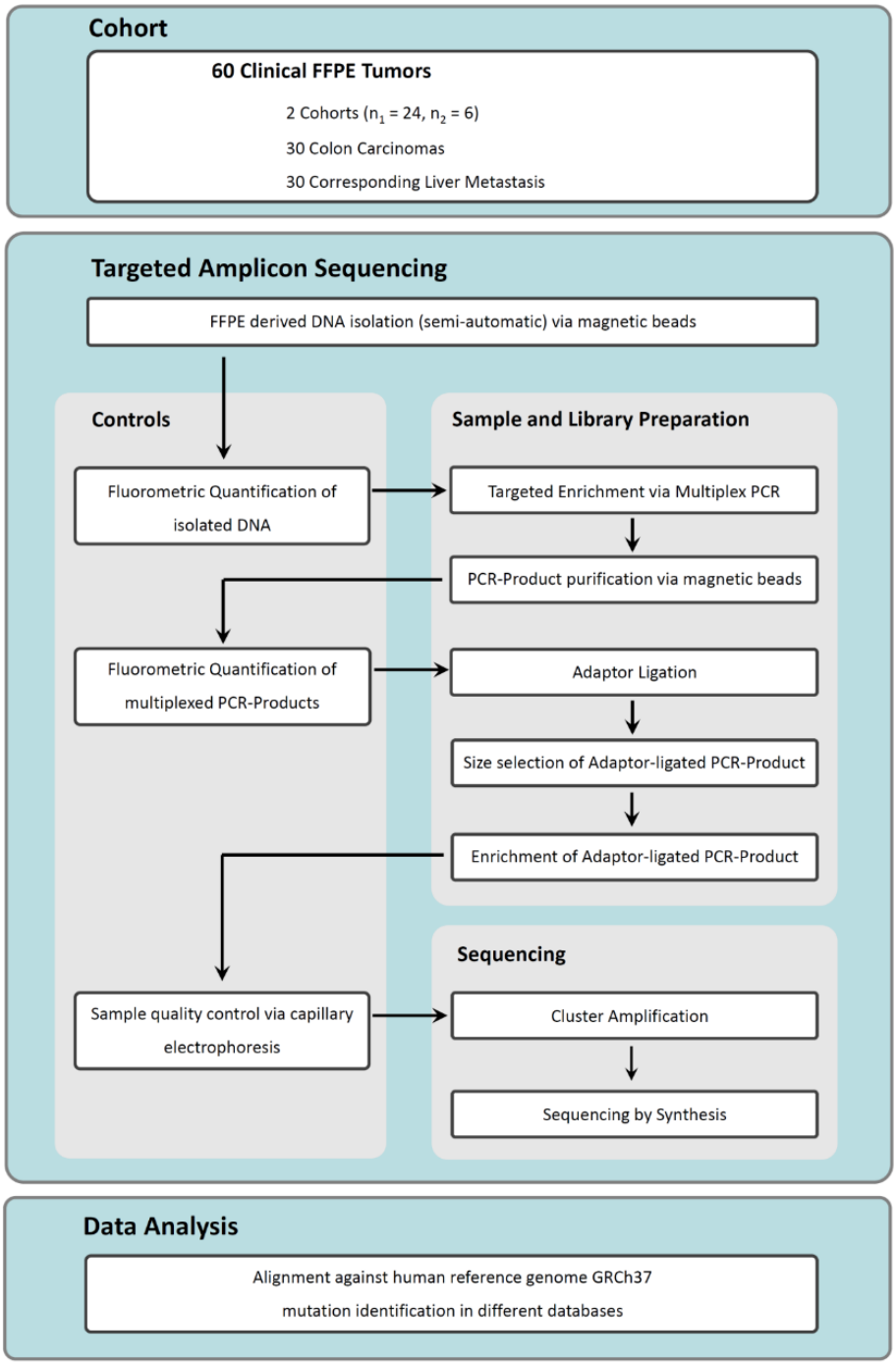

Figure 5: Methods Flowchart

3.4.1 Sample Preparation

The hematoxylin and eosin (H&E) stained samples from each subject were reviewed by an experienced pathologist. Tumor area was circled and the tumor cell percentage was evaluated and ranged between 10% to 90%. Consecutive slides were used for guided macrodissection.

3.4.2 DNA Isolation

The sample material consisted of unstained two 10 µm paraffin sections that were stored and transported at ambient temperature in low retention tubes. The genomic DNA was extracted from the samples using a semiautomatic DNA isolation kit (RSC DNA FFPE Plus Kit AX4920 custom, Promega Maxwell, Wisconsin USA) on the appertaining Maxwell RSC (Maxwell RSC Instrument AS4500, Promega Maxwell, Wisconsin USA) according to the manufacturer's manual (R29X). Proteinase K end concentration was 20 mg/ml. The proteinase K mix was incubated over-night at 70°C (Eppendorf ThermoMixer F1.5, Eppendorf, Hamburg DEU) and eluted in 50 µl nuclease free water (Plus Kit AX4920). The DNA obtained was flurometrically quantified (Qubit1 fluorometer; Invitrogen, Carlsbad, USA) using the Qubit dsDNA HS assay kit (Life Technologies, Gent, Belgium). The DNA was diluted to 45 ng DNA/18 µl nuclease free water (Plus Kit AX4920) for a 4-pool panel.

3.4.3 Targeted Enrichment via Multiplex PCR

The multiplex PCR was performed according to the manufacturer's protocol (GeneRead Targeted Panels v2, Handbook 06/15, Qiagen). For each patient sample, 4 different primer pools were selected to avoid unspecific PCR fragments. For each different primer pool, a master mix was prepared. For the multiplex PCR, the master mix for one sample contained 4.4 µl GeneRead DNA Seq panel PCR Buffer (Qiagen, Hilden DEU), 2.7 µl custom primer pool (Qiagen) including primers for the following genes: *AKT1* (NM_005163.2)*; AKT2* (NM_001626.6)*; ARID1A* (NM_006015.4)*; ARID1B* (NC_000006.11)*; APK3; ATM; BAP1; BCLAF1; BRAF (11, 15); BRCA1; BRCA2* (NM_000059.3)*; CRAF; EGFR (18-21)* (NM_201284.2)*; ERBB2 (5, 6, 15, 20, 23, 29); GNA11; GNAQ; GNAS* (NM_000516.6)*; IDH1 (4); IDH2 (4); KDM6A; KIT (9-11, 13, 17, 18); KRAS (2-4)* (NM_033360.4)*; MAP2K1; MAP2K2; MAPK1; MDM2; MET (3, 8, 11, 14, 19); MLH1* (NM_000249.4)*; MSH2* (NM_000251.3)*; NF1* (NM_000267.3)*; NRAS (2-4)* (NM_002524.5)*; PALB2*

(NM_024675.3)*; PBRM1* (NM_001350079.2)*; PDGFRa (12, 14, 18); PIK3CA (3, 5, 10, 16, 21)* (NM_006218.4)*; PTEN* (NM_000314.8)*; RNF43; RPA1; SF3B1 (14-16), SMAD4* (NM_005359.6)*; SMARCA2; SMARCA4; SMARCB1; STK11* (NM_000455.4)*; TP53* (NM_000546.6); *TSC1* (NM_001162426.2)*; TSC2* (NM_000548.5), 1.5 µl HotStarTaq DNA Polymerase (6 U/µl, Qiagen) and 9 µl DNAse-free water (Qiagen). 4 µl diluted FFPE DNA (10 ng) and 16 µl master mix were added in each well (96 well PCR plate, non-skirted black grid 4ti-0750, 4titude, Wotton GBR) and covered by PCR pate seals (PCR Seal 4ti-0500, 4titude).

The multiplex qPCR was performed according to the manufacturer's manual (Labcycler, SensoQuest Bioké, Göttingen DEU). The PCR amplification was performed using following PCR conditions: Initial Denaturation (95°C; 15 min; 1 cycle); Denaturation (95°C; 15 sec; 23 cycles); Annealing (60°C; 4 min; 23 cycles) Elongation (72°C; 10 min; 23 cycles); Hold (4°C).

3.4.4 DNA purification

PCR products from each patient were pooled to a total volume of ~80 µl. 25 µl PCR product was transferred into a PCR Strip (PCR KOMBI low profile 8-Strips #621930, Biozym, Oldendorf DEU). 45 µl (1.8x) AMPure XP Beads (Inc. #A63881, Beckman Coulter, Krefeld DEU) were resuspended in each sample and were incubated for 5 min at RT. The PCR products were placed into a magnet plate (DynaMag™-96 Side Magnet #12331D, Thermo Fischer Scientific, Massachusetts USA) for 5 min at RT and the supernatant was discarded. 170 µl ethanol (Ethanol 80 %, #32205 Riedel-de Haën Honeywell, Seelze DEU) was added to the sample, incubated for 30 sec and the supernatant was discarded. The beads were dried for 7 min at RT. The samples were resuspended in 28 µl dH_2O (Nuclease free water- AQUA AD injectabilia, B. Braun, Melsungen DEU) and were incubated for 2 min at RT. Subsequently, the samples were incubated for 5 min at RT in the magnet plate. 26 µl supernatant was transferred into PCR strips and the strips were stored at -20°C until further use. The DNA concentration was determined fluorometrically using the Qubit1 dsDNA HS assay kit (Life Technologies).

3.4.5 Library Preparation- Adaptor Ligation

The following steps were performed in the Tecan Freedom Evo 150 (Tecan Group AG, Männedorf, CHE). ~300 ng purified multiplexed PCR product in 55.5 µl from each sample were mixed with 3 µl End Prep Enzyme Mix and 6.5 µl End Repair Reaction Buffer [10x) (NEBNext Ultra DNA Library Prep Kit for Illumina, New England BioLabs, Ipswich USA) and mixed thoroughly in a nuclease free tube (DNA LoBind Tube 1.5 mL, Eppendorf, Hamburg DEU). The samples were incubated in a thermocycler (On Deck Thermal Cycler-ODTC 96 (INHECO, Planegg, DEU)) with heated lid set to 75°C for 30 min at 20°C, afterwards for 30 min (65°C) and then at 4°C. For 300 ng purified multiplexed PCR product (Multiplex Oligos for Illumina (NEB #E7335, #E7500) (NEBNext Kit) 15 µl Blunt/TA Ligase Master Mix (NEBNext Kit), 2.5 µl NEBNext Adaptor for Illumina (1:10) and 1 µl Ligation Enhancer (NEBNext Kit) were added to the End Prep reaction mixture and were mixed thoroughly (Total volume-83.5 µl). The sample was incubated at 20°C for 15 min (ODTC 96) and 3 µl of USER^{TM} Enzyme (Multiplex Oligos for Illumina (NEB #E7335, #E7500) was added. The sample was incubated at 37°C for 15 min with heated lid set to 47°C (ODTC 96).

3.4.6 Size Selection of Adaptor-ligated DNA

Size Selection of Adaptor-ligated DNA was performed according to the manufacturer's protocol (NEBnext for Illumina) with following adaptations: For size selection of adaptor-ligated DNA, 150 bp were chosen as approximate insert size (Total Library Size-270 bp; 1st Bead Selection 65 µl Beads; 2nd Bead Selection 25 µl Beads). 13.5 µl of dH_2O (Nuclease free water- AQUA AD injectabilia, B. Braun, Melsungen DEU) were added to the ligation reaction (100 µl total volume). 65 µl ((0.65X) 1st Bead Selection) of resuspended AMPure XP Beads (Inc. #A63881, Beckman Coulter, Krefeld DEU) were added to the ligation reaction and were mixed well. The samples were incubated for 5 min at RT. The plate (Hard-Shell® 96-Well PCR Plates, #hsp9631, BIO-RAD, Hercules, USA) was placed on a magnetic stand (96S Super Magnet Plate (A001322), Alpaqua, Massachusetts USA) to separate the beads from the supernatant. The samples were incubated for 5 min at RT and the supernatant containing the DNA was transferred to a new tube. The beads containing the large fragments were discarded. 25 µl ((0.25X) 2nd Bead Selection) resuspended AMPure XP Beads were added to the supernatant and were mixed. The samples were

incubated for 5 min at RT. The plate (#hsp9631) was placed on a magnetic stand and after 5 min, the supernatant that contained unwanted DNA was discarded. 150 µl of ethanol (Ethanol 80 %, #32205) was added to the plate while on a magnetic stand. The samples were incubated at RT for 30 sec and subsequently the supernatant was discarded. The beads were air dried for up to 5 min while the plate was on the magnetic stand with the lid open. The target DNA was eluted from the beads into 17 µl dH_20. The samples were incubated for 2 min at RT. The plate was placed on the magnetic stand and the solution was incubated for 5 min. 15 µl were transferred into a new plate.

3.4.7 Enrichment of Adaptor-ligated PCR-Product

25 µl NEBNext Q5 Hot Start HiFi PCR Master Mix (NEBNext Ultra Kit) and 5 µl Index Primer (Multiplex Oligos for Illumina NEB #E7335, #E7500, Universal PCR Primer (NEBNext Ultra Kit) were added to 15 µl Adaptor Ligated DNA Fragments to a total volume of 50 µl, mixed thoroughly and spun down (Mini-centrifuge 3-1810; neoLab, Heidelberg DEU). The plate was placed on a thermocycler and PCR amplification was performed using following PCR conditions: Initial Denaturation (98°C; 30 sec; 1 cycles); Denaturation (98°C; 10 sec; 8 cycles); Annealing/Extension (65°C; 75 sec; 8°cycles) Final Extension (65°C; 5 min; 1 cycles); Hold (4°C). The cleanup of the PCR Reaction was performed according to the manufacturer's protocol (NEBnext for Illumina, NGS Sample Preparation, New England BioLabs) The size distribution and quality were determined by capillary gel electrophoresis (4200 TapeStation, Agilent Technologies, Santa Clara, USA) using the D1000 ScreenTape Assay according to the manufacturer's protocol. As sample buffer and DNA ladder, the D1000 sample buffer and the D1000 ladder were used.

3.4.8 Cluster Amplification and Sequencing by Synthesis

The prefilled reagent cartridge (MiSeq Reagent Kits 300 cycles v2 MS-102-2002, Illumina, San Diego USA) was prepared according the Illumina MiSeq System guide #1000000061014 with the following adaptations: 5µl of 2nM library were denatured with 5µl 0.2 nM NaOH, spun down (Mini-centrifuge 3-1810; neoLab, Heidelberg DEU) and incubated 5 min at RT. 10 µl denatured library were diluted in 990 µl Hybridization Buffer HT1 (MiSeq Kit). 2 µl 10 nM PhiX (PhiX Control v3, Illumina), 3 µl TE/0.05% Tween (Tween20, P9416-100 mL, Sigma-Aldrich, St. Louis USA) and 5µl 0.2 nM NaOH (Sodium

hydroxide solution BioUltra, 10 M in H2O #72068, Sigma-Aldrich) were spun down and were incubated for 5 min at RT. 10 µl PhiX were diluted in 990 µl HT1. PhiX/HT1 was further diluted to 12.5 pM. The library and the PhiX mix were mixed to a final volume of 612 µl containing 480 µl 20 pM denatured DNA, 120 µl pre-chilled HT1 and 12 µl 12.5 pM PhiX. Afterwards, the library mix was loaded onto the reagent cartridge and the run using MiSeq Control Software was set up. The flow cell was cleaned and dried thoroughly. The wash buffer PR2 (MiSeq Kit), the reagent cartridge and the flow cell were loaded on the MiSeq (Illumina) and the run was started.

3.4.9 Data Analysis

The raw NGS data were mapped against the human reference genome GRCh37 (human hg 19). Sequence alignment was performed on the Biomedical genomic workbench 5.0.1 (Qiagen Bioinformatics) transforming raw NGS data into relevant findings. For visualization and interactive exploration of the data, the Integrative Genomics Viewer (IGV) (Broad Institute, Massachusetts USA) was used. Variants were annotated using the following databases: Single Nucleotide Polymorphism Database (dbSNP) (Sherry et al., 2001); HapMap (International HapMap et al., 2010); 1000 genome data (Genomes Project et al., 2010). Genome Annotation was performed using the COSMIC (Tate et al., 2018), ClinVar (Landrum et al., 2018) and OnkoKB database (Chakravarty et al., 2017). Non annotated mutations were included in the study. Following mutations were excluded: mutations with an allele frequency < 2%, < 50 unique reads, < 5 counts or a read length < 20 nt.

3.5 Statistics – Cohort A and B

Genes were divided into four groups according to their role in CRC: proliferation, genomic instability, insensitivity to anti-growth signals, evade apoptosis. One gene can belong in more than one group (Figure 4). It was tested whether one gene has a statistically significant association to the linear or to the parallel model, to tumor localization, to histotype, to resection time or necrosis level. Only mutations with an allele frequency > 10% and < 50 unique reads were counted and statistically analyzed.

For statistical and graphical analyses, the R statistical programming environment (v3.6.2) was used. The Shapiro–Wilks test was applied to test for gaussian distribution of the data. For dichotomous variables, either the Wilcoxon Mann–Whitney rank sum test for non-parametric variables or two-sided Student's t-test for parametric variables was used. For ordinal variables with more than two groups, either the Kruskal–Wallis test for non-parametric variables or the ANOVA for parametric variables was used to detect group differences. The level of statistical significance was defined as $p \leq 0.05$.

4. Results

4.1 Overview of identified mutation variants

4.1.1 Sequencing performance cohort A

Cohort A includes the patients whose data were collected and analyzed in this project.

132 mutations were identified from 30 samples, 61 in the primary tumor and 71 in the corresponding liver metastasis. One sample had no mutations in any of the analyzed gene regions.

The average base coverage of all sequenced genes was 989.2 reads ±1261.8. In total eleven genes from all samples had a suboptimal coverage between 54 and 100 reads, 27 genes had a coverage of 101-300 reads, 29 genes had a coverage of 301-600 reads and 65 genes had a high coverage of 625-7474 reads. The average amount of mapped bases of all samples was 97.5% ±0.67 and the average DNA concentration per sample after the first purification was 74.4 ng/µl ±36.02.

The number of mutations per tumor ranged from 1 to 8 (mean 2.2). In the majority of tumors (N=38/60) only one or two mutations were detected. The most frequently mutated genes were *TP53* (60% in primary; 70% in metastasis) and *KRAS* (33% in primary; 43% in metastasis) (Table 4; Figure 6 A, B). 31% of the mutations showed a low (3%-10%), 29% a medium (10%-30%), 32% a high (30%-60%) and 8% very high (60%-90%) allele frequency.

Mutated with 10%-90% allele frequency were the following genes: *KRAS (N=22); ARID1A (N=4); ARID1B (N=2); BRCA2 (N=3); NRAS (N=2); PBRM1 (N=1); PIK3CA (N=6); PTEN (N=2); SMAD4 (N=5); TP53 (N=37); TSC2 (N=5).* Following mutated genes were identified as unique in the primary CRC tumors: *ARID1B; SMAD4; BRCA2; NF1; TSC1; TSC2; AKT2; MAPK3; MSH2; MLH1; PALB2; ARID1A; PIK3CA; KRAS, TP53* (Table 6).

Table 6: Summary of the mutations identified in primary CRC of Cohort A. Total mutations primary tumor N=61; corresponding liver metastasis total mutations (N=71); mutations unique in primary CRC (N=26); mutations unique in metastasis (N=32).

Gene	Numer of Mutations			
	Primary (Total)	Metastasis (Total)	Unique in primary	Unique in metastasis
ARID1B	7	5	6	4
SMAD4	4	3	3	2
BRCA2	3	1	2	0
NF1	2	1	2	1
TSC1	2	2	2	2
TSC2	4	6	2	4
AKT2	1	0	1	0
MAPK3	0	1	1	0
MSH2	1	0	1	0
MLH1	0	1	1	0
PALB2	1	0	1	0
ARID1A	2	2	1	1
PIK3CA	3	5	1	3
KRAS	10	13	1	4
TP53	18	21	1	4
NRAS	1	1	0	0
PBRM1	1	1	0	0
AKT1	0	1	0	1
EGFR	0	1	0	1
PTEN	1	2	0	1
GNAS	0	2	0	2
STK11	0	2	0	2
Σ	**61**	**71**	**26**	**32**

From the mutated genes identified as unique in primary, the following had an allele frequency above 10%: *ARID1A* (N=2); *BRCA2* (N=1); *KRAS* (N=3), *PTEN* (N=1); *TP53* (N=2); *SMAD4* (N=2). For two mutations identified as unique in primary tumor (*BRCA2* (N=1); *KRAS* (N=1)), the respective genes were not identified in the metastasis due to missing reads. Data are illustrated in Figure 6A (primary) and 6B (metastasis). Difference in pairs was calculated by the subtraction of the mutations unique in liver metastasis from the mutations unique in primary CRC (Figure 6C). Following genes were more often

mutated in primary CRC: *ARID1B, BRCA2, SMAD4, NF1, AKT2, MAPK3, MSH2, MLH1 and PALB2*. Following genes were more often mutated in the liver metastasis: *AKT1, EGFR, PTEN, TSC2, PIK3CA, GNAS, STK11, KRAS and TP53*. *TSC1, ARID1A, NRAS* and *PBRM1* showed no difference in pairs.

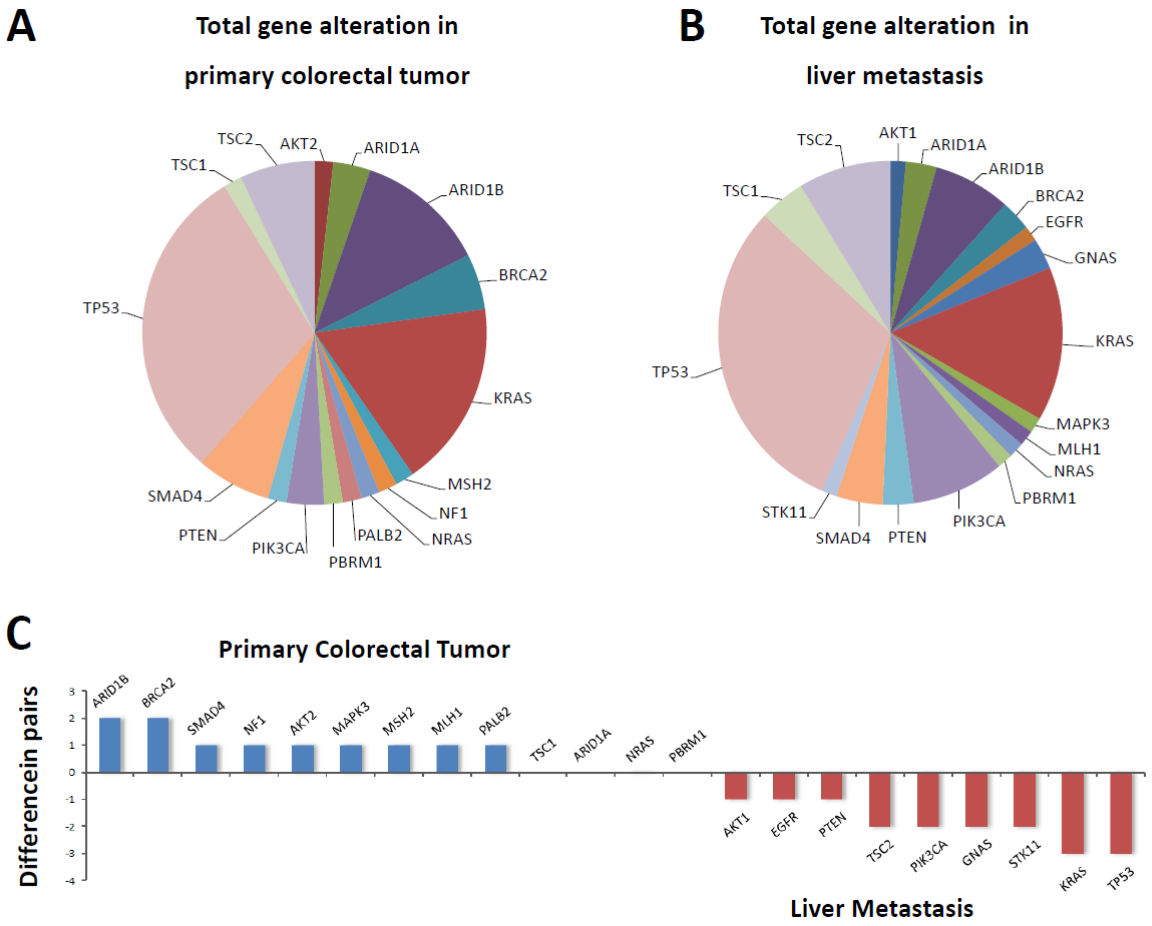

Figure 6: Summary of the mutations identified in primary CRC (A) and liver metastasis (B) that can influence tumor growth and dissemination (N = 57; N = 69 respectively). C Difference in pairs calculated as the difference between the mutations unique in primary CRC and mutations unique in liver metastasis.

4.1.2 Mutation analysis Cohort A

The association of different parameters such as histotype, resection time, localization and necrosis status with certain gene mutations was statistically analyzed. No significant association was observed between localization or resection time in any of the genes. Two gene mutations in *PTEN* and *TP53* in the primary tumor were significantly associated with the usual adenocarcinoma histotype (p=0.0046; p=0.036 respectively). One gene, *MLH1* in metastasis was significantly associated with the mucinous histotype (p=0.0142). *PIK3CA* was significantly associated with necrosis level (p=0.05).

Next, it was investigated whether certain gene mutations are more prominent in either the linear or the parallel model. *SMAD4*, *ARID1A* and *PTEN* mutations were significantly more frequent in tumors of the parallel model rather than in tumors of the linear model (Figure 7).

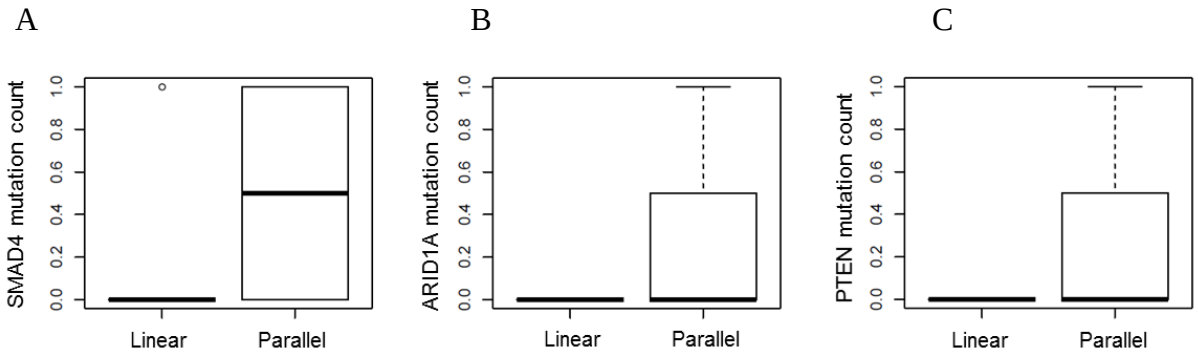

Figure 7: Primary tumor mutation count in the linear and the parallel model of cohort A for *SMAD4* (A), *ARID1A* (B) and *PTEN* (C). *SMAD4*, *ARID1A* and *PTEN* are identified as mutated in the parallel model (p=0.0058; p= 0.014; p= 0.014, respectively)

4.1.3 Mutation analysis Cohort B (cBioPortal)

In order to validate the results of this project (cohort A), a second cohort (cohort B) was designed using data from Brannon (Brannon et al., 2014). The colorectal adenocarcinoma data of 66 patients and the corresponding liver metastasis data were retrieved from cBioPortal (MSKCC, Genome Biol 2014).

Out of 66 patients, 7 had unique gene mutations in the primary tumor and thus can be allocated to the parallel model of progression.

26 genes were identified as mutated, of which seven (*BAP1; KDM6A; PBRM1; PDGFRB; PIK3CA; STK11; TP53*) were unique in the primary tumor and eight (*ATM; BRCA2; EGFR; MAP2K1; NF1; PIK3CA; PTEN; SMAD4*) were unique in the liver metastasis (Figure 7; Figure 8). Following sequenced genes had no mutations: *AKT2; ARID1A; ARID1B; CDKN2A; FGFR1; FGFR2; GNA11; GNAQ; IDH1; IDH2; KIT; MAPK1; MAPK3; MAPK8; MDM2; MLH1; MSH2; MYC; NF2; PALB2; PRKAR1A; RAF1; RB1; RET; RNF43; RPA1; SMARCA2; SMARCA4; SMARCB1; TP73; TSC2* (data not shown).

The number of mutations per tumor in the selected genes ranged from 1 to 4 (mean 2.1). In the majority of tumors (N=78/132) only one or two mutations were detected (Figure 9 B). The most frequently mutated genes were *TP53* (89% in primary; 87% in metastasis) and *KRAS* (56% in primary and metastasis).

The gene mutations detected in cohort B were not significantly associated neither with the tumor localization nor with the tumor histotype or with the necrosis level. Four primary tumor gene mutations (*KDM6A, MSH6, STK11, TSC1*) and two metastatic tumor mutations (*MSH6; TSC1)* that were only detected once, were significantly associated (p= 0.0043) with the progression models (Table 7). One gene, *MAP2K1,* was significantly associated with the resection time (p=0.0126).

Table 7: Summary of the mutations identified in primary CRC total (N=148); corresponding liver metastasis total (N=148); mutations unique in primary CRC and metastasis (N=8).

Gene	Number of Mutations			
	Primary (Total)	Metastasis (Total)	Unique in primary	Unique in metastasis
AKT1	1	1	0	0
ATM	1	2	0	1
BAP1	2	1	1	0
BRAF	3	3	0	0
BRCA1	1	1	0	0
BRCA2	0	1	0	1
EGFR	0	1	0	1
FBXW7	7	7	0	0
FGFR3	1	1	0	0
GNAS	2	2	0	0
KDM6A	1	0	1	0
KRAS	36	36	0	0
MAP2K1	1	2	0	1
MET	1	1	0	0
MSH6	1	1	0	0
NF1	1	2	0	1
NRAS	2	2	0	0
PBRM1	2	1	1	0
PDGFRB	2	1	1	0
PIK3CA	13	12	2	1
PTEN	1	2	0	1
SMAD4	8	9	0	1
STK11	1	0	1	0
TP53	57	56	1	0
TSC1	1	1	0	0
TSHR	2	2	0	0
Σ	**148**	**148**	**8**	**8**

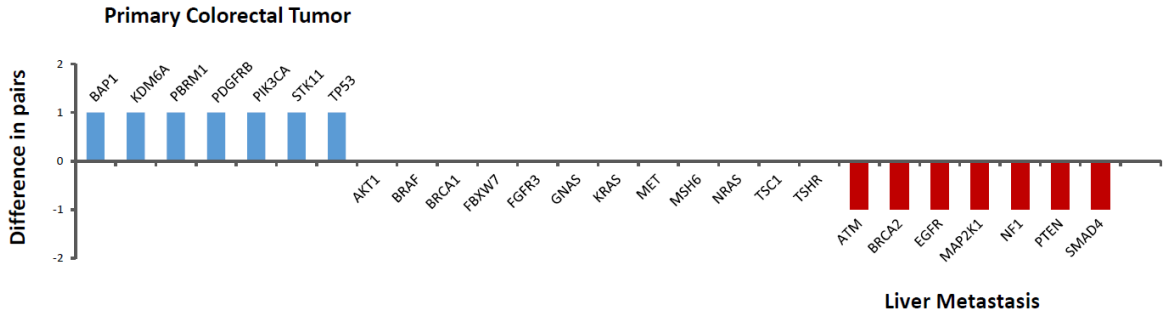

Figure 8: Difference in pairs of cohort B, calculated as the difference between the mutations unique in primary CRC and mutations unique in liver metastasis.

4.1.4 Increased mutation rate in the parallel model

To evaluate whether there is a relationship between the mutation rate and the tumor model, data from both cohort A and cohort B were analyzed. It was shown that in both cohorts, in the linear model there are significantly less primary tumor mutations than in the parallel model ($p = 0.0695$; $p = 0.0366$ for cohort A and B respectively) (Figure 9).

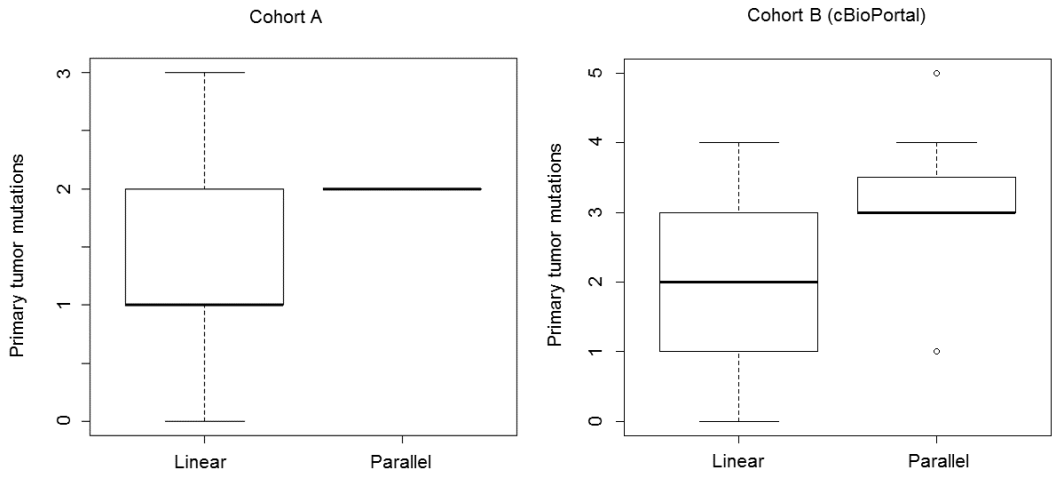

Figure 9: Mutation count in the primary tumor in the linear and parallel model for cohort A (left) and cohort B (right). In both cohort A and cohort B the linear model shows significantly less primary tumor mutations than the parallel model ($p = 0.0695$; $p = 0.0366$ for cohort A and B respectively). The Wilcoxon Mann–Whitney rank sum test was applied.

4.2 Cases referring to parallel progression

4.2.1 Patient 17

A 66 years old patient was diagnosed with CRC in the ascending colon and a synchronous liver metastasis in liver segment I (TNM STAGE: T3G2pN1a pM1). Necrosis was absent in the primary tumor and low in the metastasis (Figure 10).

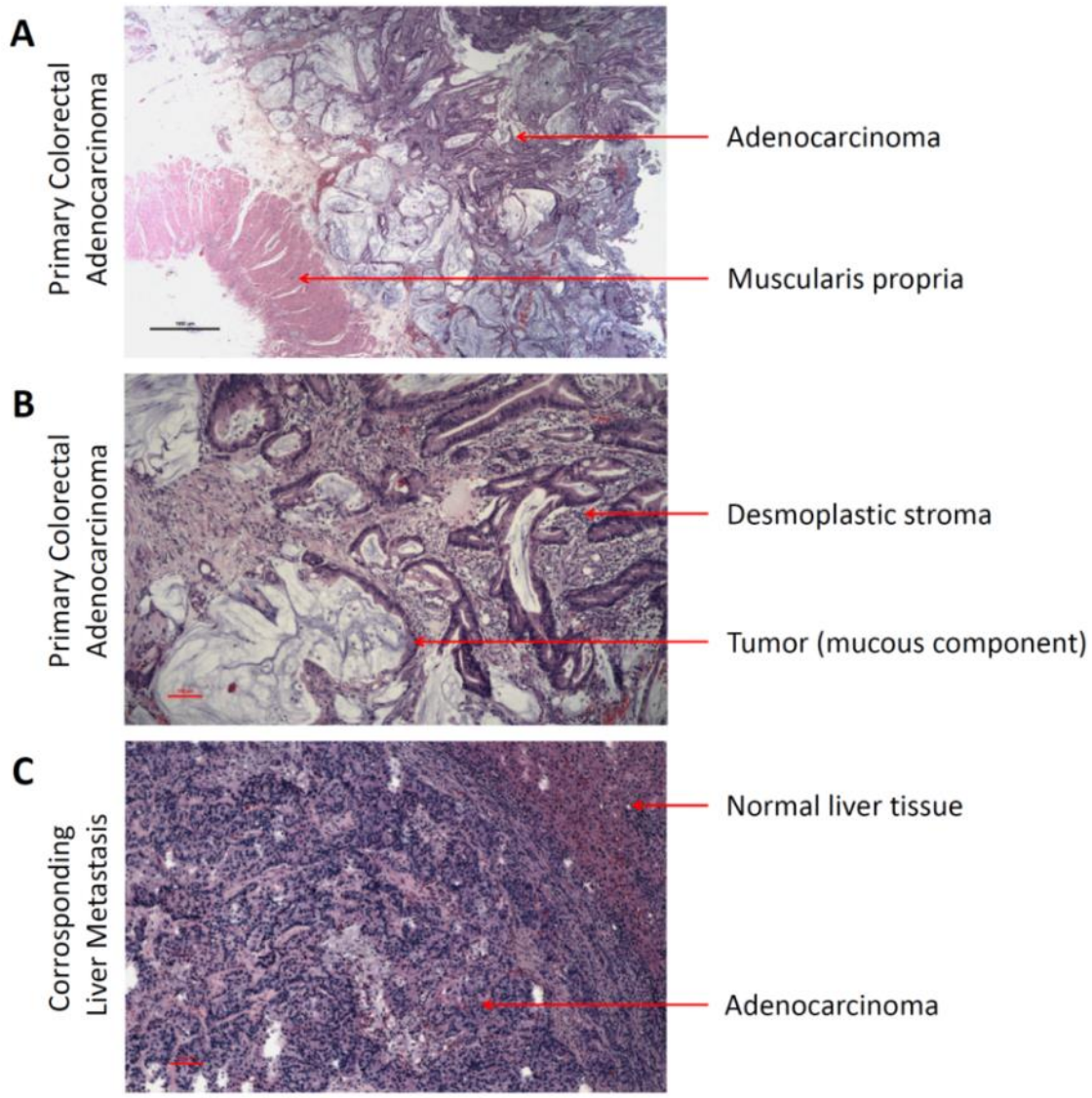

Figure 10: H&E stained slides of colorectal carcinoma and corresponding liver metastasis. (A) Primary Colorectal Adenocarcinoma (20x magnification): Clusters of atypical columnar-cell proliferations with partial extracellular mucinous differentiation and slight desmoplastic stroma. Tumor cells infiltrating the submucosa. No tumor-infiltration of the muscularis propria. (B) Primary Colorectal Adenocarcinoma (200x magnification): Clusters of atypical columnar- and goblet-cells with moderate nuclear atypia. Tumor formations showing partial extracellular mucinous

differentiation (“mucous-lakes”). Tumor cells are embedded in desmoplastic stroma. Furthermore, an accompanying increased inflammatory reaction can be observed. (C) Corresponding liver metastasis (magnification 200x): Sheets of solid atypical glandular epithelial cell proliferations with moderate nuclear atypia. A partial slight desmoplastic stroma reaction as well as an increased inflammatory reaction can be observed. Normal liver tissue is marginally found.

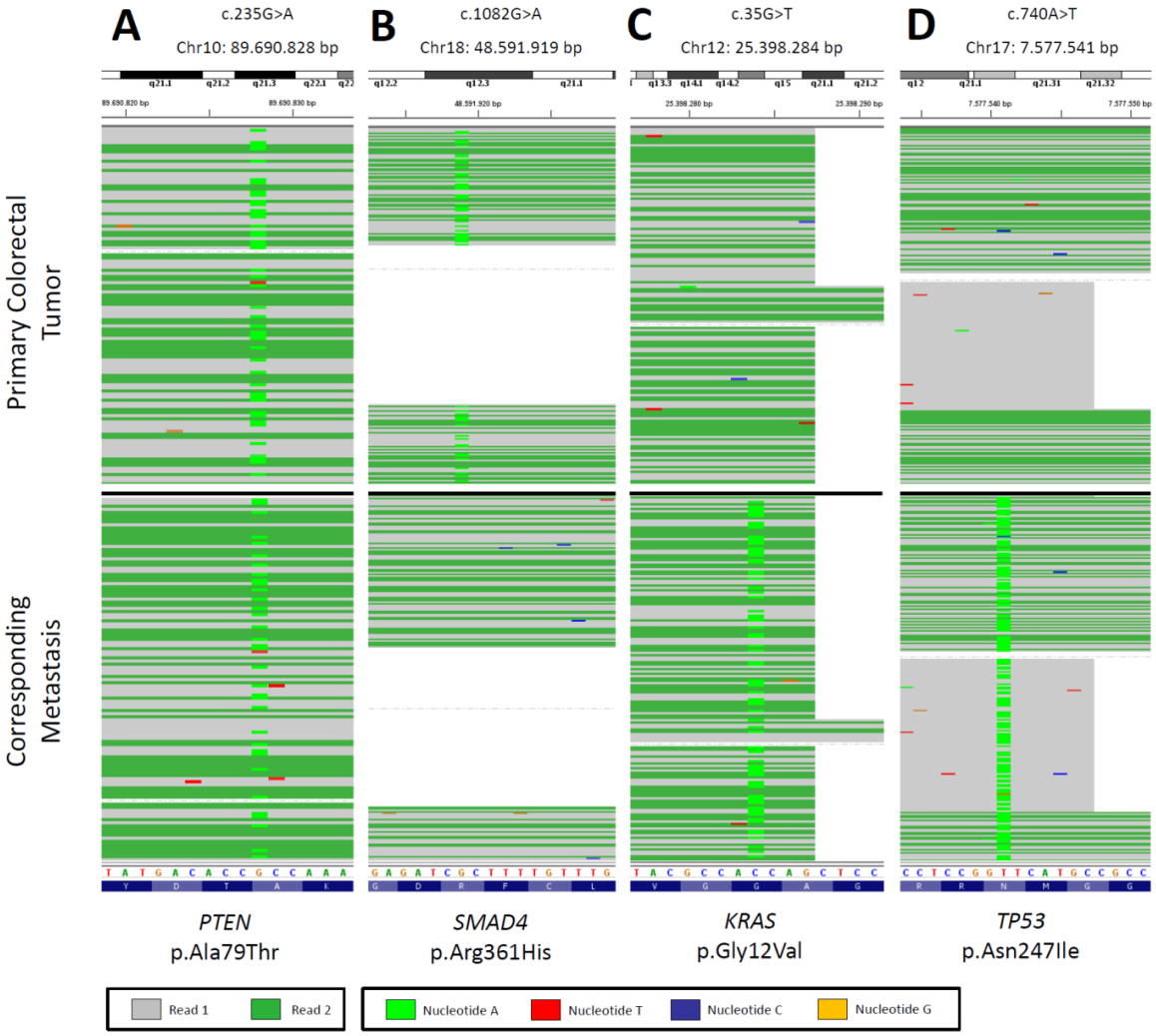

Figure 11: Exemplary raw NGS data from primary colorectal tumor (upper part) and corresponding liver metastasis (lower part) showing 4 different mutation sites. Grey or dark green lines show the two reads (forward or reverse strand). Mutations are highlighted as bars (light green for nucleotide A; red T; blue C; orange G). The single-letter amino acid code is in 5' → 3' direction. The codons (triplets) are in (A) *PTEN* 5'→3'; (B) *SMAD4* 5'→3'; (C) *KRAS* 3'→5' (D) *TP53* 3'→5' direction. (A) The *PTEN* mutation (c.235G>A) can be observed in the in the primary tumor and in the metastasis. (B) The *SMAD4* mutation (c.1082G>A) can only be observed in the primary tumor. (C) The *KRAS* mutation (c.35G>T) and the (D) *TP53* mutation (c.740A>T) can only be observed in the metastasis. As reference genome, GRCh37 was used.

A *PTEN* p.Ala79Thr (c.235G>A) 10q23.3 mutation was identified both in the primary and in the metastasis with an allele frequency of 47 % and 789 reads in the primary and an allele frequency of 34 % and 524 reads in the metastasis, respectively (Figure 11 A). The mutation is listed in the ClinVar database and is likely to be benign. (ID: 41682: SCV000840467.1). A *SMAD4* p.Arg361His (c.1082G>A) 18q21.2 mutation was identified as unique in primary tumor with an allele frequency of 42% and 325 reads (Figure 11 B). The mutation is listed in the ClinVar database and described as pathogenic in neoplasm of the large intestine (ID: 24832: SCV000504619.1). A *KRAS* p.Gly12Val (c.35G>T) 12p12.1 mutation was identified unique in metastasis with an allele frequency of 46% and 864 reads (Figure 11 C). The somatic mutation is listed in the ClinVar database (ID: 12583: SCV000504474.1) as pathogenic in neoplasm of the large intestine. A *TP53* p.Asn247Ile (c.740A>T) 17p13.1 mutation was identified as unique in metastasis with an allele frequency of 77% and a coverage of 2055 unique reads (Figure 11 D). The mutation is listed in the ClinVar database described with uncertain significance in hereditary cancer-predisposing syndrome (ID:184882: SCV000214839.3).

4.2.2 Patient 9

56 years old male patient was diagnosed with CRC in the rectum and synchronous liver metastasis in liver segments IV, V, VIII (initial TNM STAGE: ypT3N0). The primary tumor was delineated as a usual adenocarcinoma, while in the metastasis mucinous differentiation was present. These two morphologically different tumors are a further indication for the parallel progression model.

An insertion in *ARID1A* (p.Gln1334_Arg1335insGln; (c.3977_3978insGCA)1p36.11) was identified in both the primary and the metastasis. Namely, with an allele frequency of 60% and 1359 reads in the primary and an allele frequency of 25% and 1320 reads in the metastasis (Figure 12 A). The mutation is listed in the ClinVar database and is likely to be benign. (ID: 235338: SCV000280832.1). A *TP53* p.Arg273Cys (c.817C>T) 17p13 mutation was identified both in the primary and in the metastasis, with an allele frequency of 35% and 1509 reads in the primary, and an allele frequency of 54% and 2744 reads in the metastasis (Figure 12 B). The mutation is listed in ClinVar database with conflicting interpretations of pathogenicity (Likely pathogenic (N=4); Pathogenic (N=7); Uncertain significance (N=1)). The mutation is associated with hepatocellular carcinoma (ID: 43594: SCV000504686.1).

Novel *ARID1A* p.Gln1655* (c.4963C>T) and *ARID1B* p.His96del (c.270_272delCCA) mutations were identified as unique in the primary with an allele frequency of 21% and 781 reads (Figure 12C) and 5.7% and 104 reads respectively.

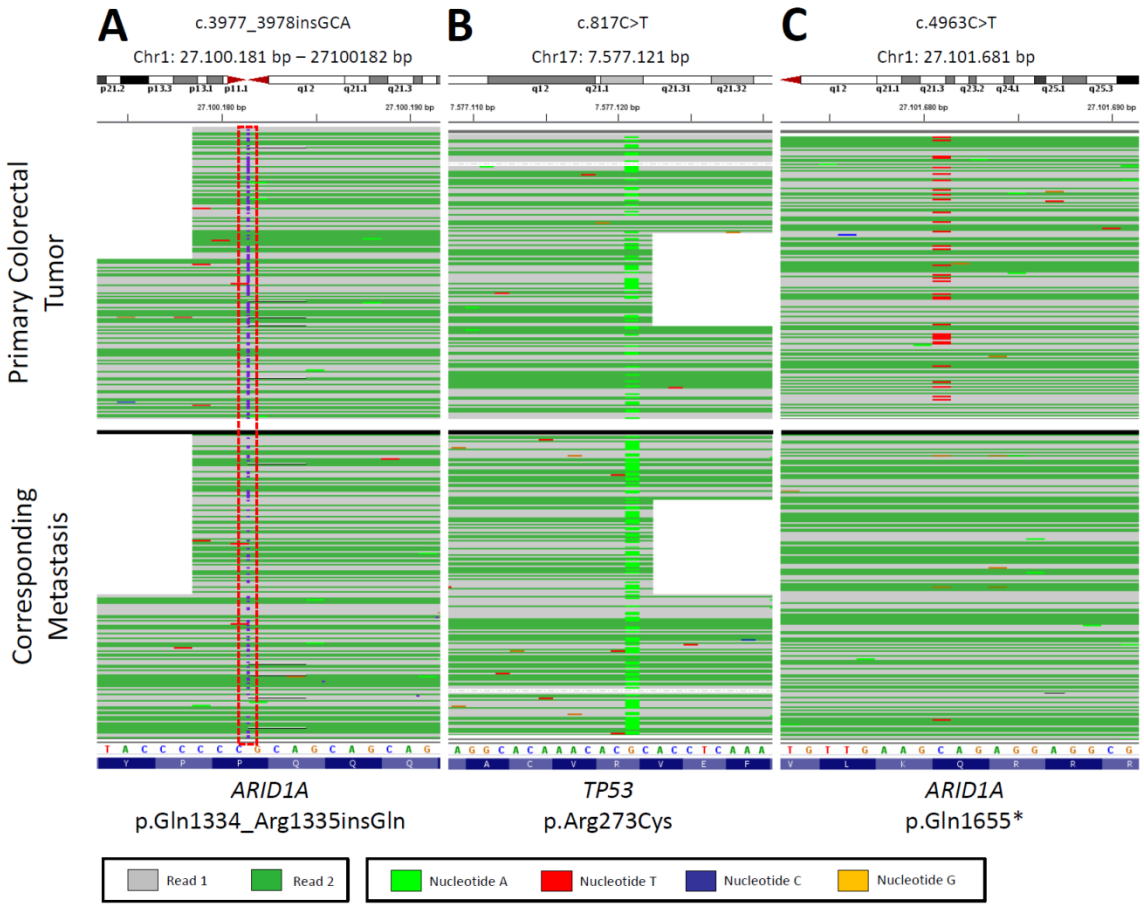

Figure 12: Exemplary raw NGS data from primary colorectal tumor (upper part) and corresponding liver metastasis (lower part) showing 3 different mutation sites. Grey or dark green lines show the two reads (forward or reverse strand). Mutations are highlighted as bars (light green for nucleotide A; red T; blue C; orange G). The insertion is depicted in blue dots (framed in red). The single-letter amino acid code is in 5' → 3' direction. The codons (triplets) are in (A) and (C) *ARID1A* 5'→3'; B) *TP53* 3'→5' direction. (A) The Gln-insertion in *ARID1A* (c.3977_3978insGCA) and the (B) *TP53* mutation (c.817C>T) can be observed in the primary tumor as well as in the metastasis. (C) The *ARID1A* mutation (c.4963C>T) can only be observed in the primary tumor. As reference genome, GRCh37 was used.

4.2.3 Patient 5

66 years old male patient was diagnosed with CRC in the ascending colon. After 1 year, a metachronous liver metastasis in liver segment III (initial TNM STAGE: pT2aN0) was operated and removed. The patient received FOLFOX and FOLFIRI as adjuvant therapy.

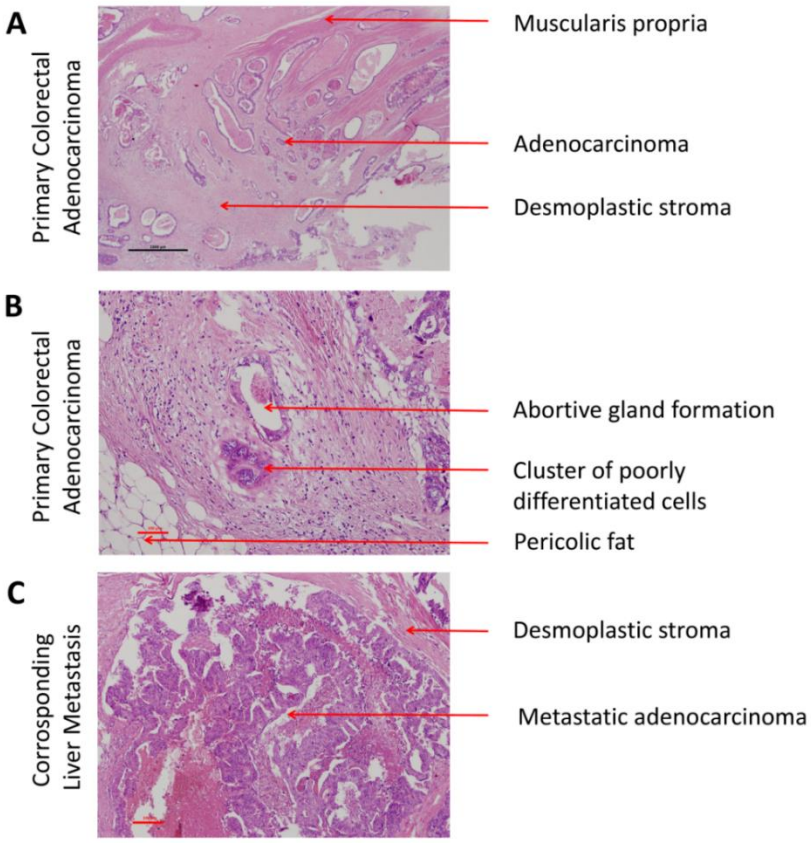

Figure 13: H&E stained slides of colorectal carcinoma and corresponding liver metastasis. (A) Primary Colorectal Adenocarcinoma (100x magnification): tumor cells form abortive glands within dense desmoplastic stroma. Invasion of muscularis propria is present. (B) Primary Colorectal Adenocarcinoma (200x magnification): At the invasive margin both moderately differentiated cells forming abortive glands and clusters of poorly differentiated cells with vesicular nuclei are present. Invasion of pericolic fat is seen. Lymphocytic infiltration in the desmoplastic stroma is mild. (C) Corresponding liver metastasis (magnification 100x): Columnar tumor cells form abortive glands. Desmoplastic stroma with mild lymphocytic infiltration is seen.

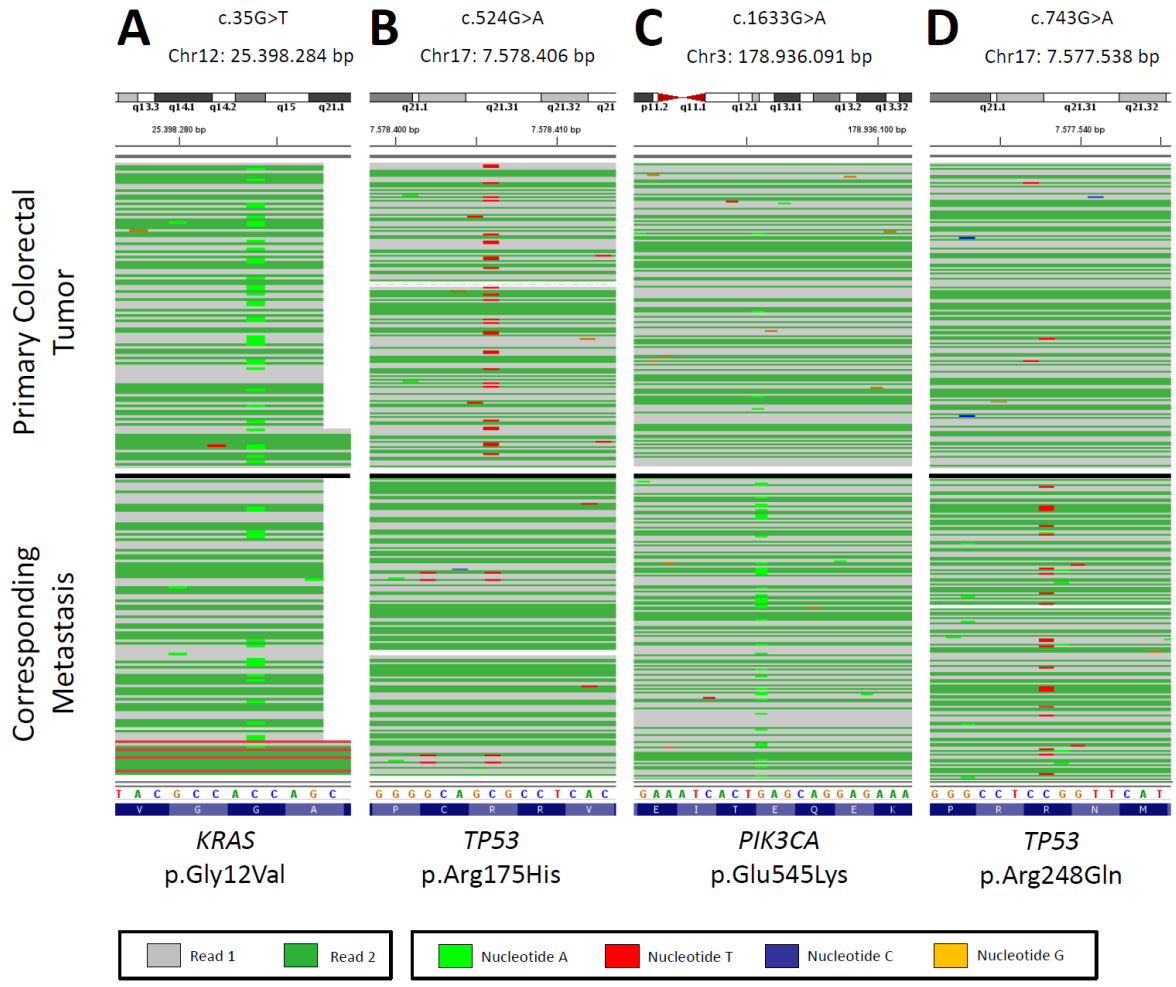

Figure 14: Exemplary raw NGS data from primary colorectal tumor (upper part) and corresponding liver metastasis (lower part) showing 4 different mutation sites. Grey or dark green lines show the two reads (forward or reverse strand). Red lines indicate an inferred insert size that is larger than expected (deletion). Mutations are highlighted as bars (light green for nucleotide A; red T; blue C; orange G). The single-letter amino acid code is in 5'→3'. The codons (triplets) are in (A) *KRAS* 3'→5'; (B) and (D) *TP53* 3'→5' (C) *PIK3CA* 5'→3' direction. (A) The *KRAS* mutation (c.35G>T) can be observed in the in the primary tumor and in the metastasis. (B) The *TP53* mutation (c.524G>A) can only be observed in the primary tumor. (C) The *PIK3CA* mutation (c.1633G>A) as well as the (D) *TP53* mutation (c.743G>A) can only be observed in the metastasis. As reference genome, GRCh37 was used.

A *KRAS* p.Gly12Val (c.35G>T) 12p12.1 mutation was identified both in the primary and in the metastasis, with an allele frequency of 30% and 457 reads in the primary, and an allele frequency of 20% and 629 reads in the metastasis (Figure 14 A). The mutation is listed in ClinVar database as pathogenic. The somatic mutation is associated with neoplasm of the

large intestine (ID: 12583: SCV000504474.1). *A TP53* p.Arg175His (c.524G>A) 17p13 mutation was identified as unique in primary tumor with an allele frequency of 13% and 521 reads (Figure 14 B). The mutation is listed in ClinVar database as pathogenic. The somatic mutation is associated with neoplasm (ID: 12374: SCV000504891.1). A *SMAD4* p.Gly419Trp (c.1255G>T) 18q48 mutation was also identified as unique in the primary an allele frequency 5.6% and 3276 reads. The mutation is listed in COSMIC database as pathogenic (ID: COSM1151580). A *PIK3CA* p.Glu545Lys (c.1633G>A) 3q26.3 mutation was identified as unique in the metastasis with an allele frequency of 21% and 1097 reads (Figure 14 C). The mutation is listed in ClinVar database as pathogenic. The somatic mutation is associated with neoplasm of the large intestine (ID:°13655: SCV000503936.1. A *TP53* p.Arg248Gln (c.743G>A) 17p13 mutation was identified as unique in metastasis with an allele frequency of 19% and 337 reads (Figure 14 D). The mutation is listed in ClinVar database as pathogenic. The somatic mutation is associated with neoplasm of the large intestine (ID: 12356: SCV000504708.1). A novel *ARID1B* mutation (p.His96del (c.270_272delCCA)) was identified as unique in primary tumor with an allele frequency of 5.7% and 246 reads.

4.2.4 Patient 29

63 years old male patient was diagnosed with CRC in the hepatic flexure with multiple synchronous liver metastases in the left lobe (initial TNM STAGE: pT3N0). The primary tumor was shown to be conventional adenocarcinoma, grade 1, while the liver metastasis exhibited mucinous differentiation in more than 95% of the tumor body. The patient received FOLFIRI and anti-EGFR therapy as adjuvant treatment.

A *KRAS* p.Lys117Asn (c.351A>T) 12p12.1 mutation was identified in both the primary and the metastasis, with an allele frequency of 30% and 1984 reads in the primary, and 12% and 1223 reads in the metastasis. The mutation is listed in ClinVar database as pathogenic and is associated with neoplasm of the large intestine (ID: 375964: SCV000504419.1). A *SMAD4* p.Arg361Cys (c.1081C>T) 18q21.2 mutation was identified in the primary with an allele frequency of 29% and 206 reads. The mutation is listed in ClinVar database as pathogenic and is associated with neoplasm of the large intestine (ID: 8543: SCV000504611.1).

4.3 Cases referring to linear progression

4.3.1 Cases developing additional mutations in the metastasis

4.3.1.1 Patient 12

75 years old male patient was diagnosed with CRC in the cecum with synchronous liver metastasis in liver segments 4b and 5 (initial TNM STAGE: pT3N2). A *KRAS* p.Gly13Asp (c.38G>A) 12p12.1 mutation was identified in the primary with an allele frequency of 63% and 851 reads and in the metastasis with an allele frequency of 26% and 898 reads (Figure 15 A). The mutation is listed in ClinVar database as pathogenic. The somatic mutation is associated with neoplasm of the large intestine (ID: 12580: SCV000504458.1).

A *SMAD4* p.Asp537Gly (c.1610A>G) 18q21.2 mutation was identified in the primary and the metastasis with an allele frequency of 55% and 849 reads in the primary and an allele frequency of 26% and 179 reads in the metastasis (Figure 15 B). The mutation is listed in ClinVar database as likely pathogenic. The somatic mutation is associated with neoplasm of the large intestine (ID: 523175: SCV000734842.1). A *TP53* p.Gly245Ser (c.733G>A) 17p13 mutation was identified in the primary and the metastasis with an allele frequency of 50% and 399 reads in the primary and an allele frequency of 20% and 632 reads in the metastasis (Figure 15 C) The mutation is listed in ClinVar database as likely pathogenic. The somatic mutation is associated with neoplasm of the large intestine (ID: 12365: SCV000504879.1). A *TSC2* p.Ala1257Val (c.3770C>T) 16p13.3 mutation was identified as unique in the metastasis with an allele frequency of 16% and 1583 reads (Figure 15 D). The mutation is listed in ClinVar database as benign. (ID: 41734: SCV000214457.4). Read 1 indicated in green was deficient (Figure 15). A *STK11* p.Pro324Leu (c.971C>T) 19p13.3 mutation was identified as unique in the metastasis with an allele frequency of 5% and 216 reads in the metastasis. The mutation is listed in ClinVar database as with uncertain significance. The mutation is associated with hereditary cancer-predisposing syndrome (ID: 135931: SCV000213929.4).

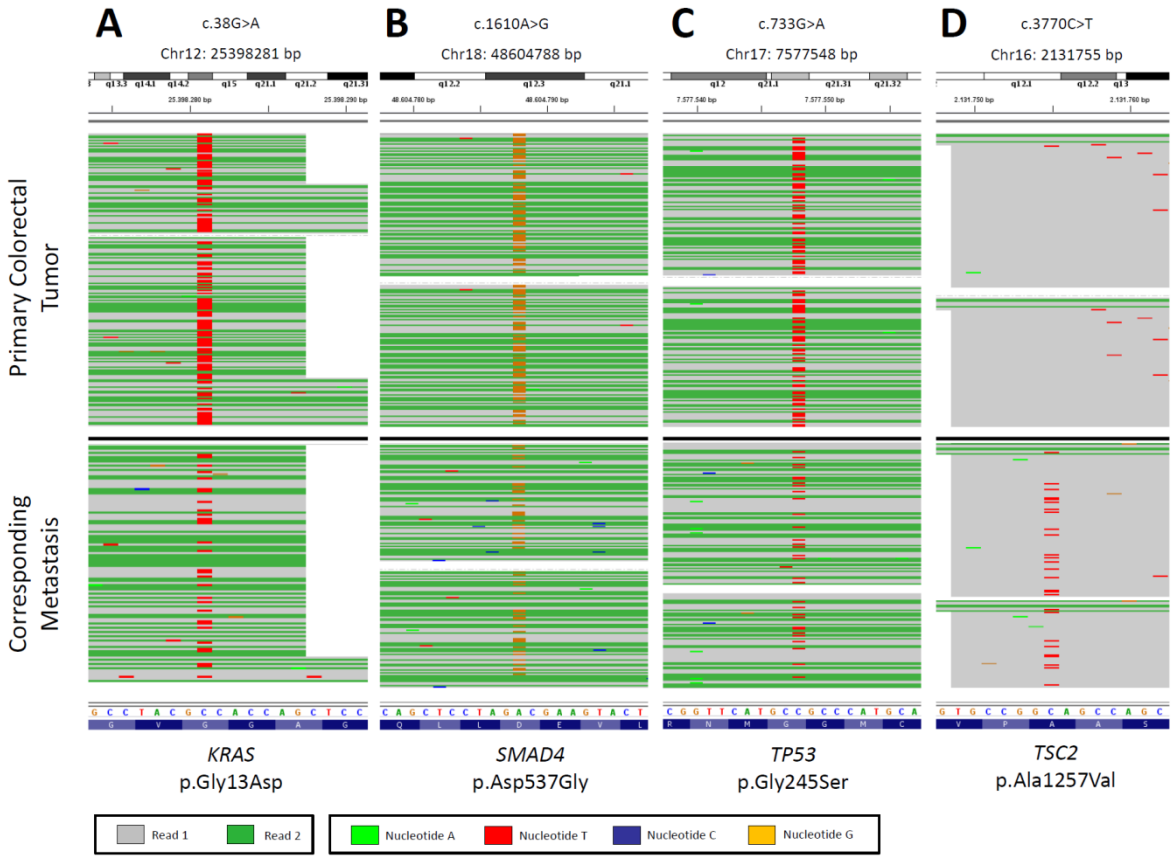

Figure 15: Exemplary raw NGS data from primary colorectal tumor (upper part) and corresponding liver metastasis (lower part) showing 4 different mutation sites. Grey or dark green lines show the two reads (forward or reverse strand). Mutations are highlighted as bars (light green for nucleotide A; red T; blue C; orange G). The single-letter amino acid code is in 5'→3'. The codons (triplets) are in (A) *KRAS* 3'→5'; (B) *SMAD4* 5'→3'; (C) *TP53* 3'→5'; (D) *TSC2* 5'→3' direction. (A) The *KRAS* mutation (c.38G>A), the (B) *SMAD4* mutation (c.1610A>G) and the (C) *TP53* mutation (c.733G>A) can be observed in the in the primary tumor and in the metastasis. (D) The *TSC2* mutation (c.3770C>T) can only be observed in the metastasis tumor. One of the strands showed poor read. As reference genome, GRCh37 was used.

4.3.1.2 Patient 4

35 years old male patient was diagnosed with CRC in the rectum. After 1 year, a metachronous liver metastasis in liver segment 4 (initial TNM STAGE: pT3N1) was operated and removed. The patient received FOLFOX, FOLFIRI and Bevacizumab as adjuvant therapy.

A *TP53* p.Arg342Ter (c.1024C>T) 17p13.1 mutation was identified both in the primary and in the metastasis with an allele frequency of 44% and 705 reads in the primary and an allele frequency of 33% and 695 reads in the metastasis. The mutation is listed in ClinVar database as pathogenic; it is associated with hereditary cancer-predisposing syndrome (ID: 182970: SCV000212808.5). An *ARID1B* p.His96del (c.270_272delCCA) 6q25.3 mutation was identified in the metastasis with an allele frequency of 5.9% and 239 reads.

4.3.1.3 Patient 13

74 years old male patient was diagnosed with CRC in the sigmoid colon. After one year, a metachronous liver metastasis in liver segments 4 and 6 (initial TNM STAGE: pT3N1b) was operated and removed. The patient received FOLFOX and Bevalizumab as adjuvant treatment.

A *TP53* p.Gly245Asp (c.734G>A) 17p13.1 mutation was identified both in the primary and in the metastasis with an allele frequency of 47% and 592 reads in the primary and an allele frequency of 44% and 558 reads in the metastasis. The mutation is listed in ClinVar database as likely pathogenic and is associated with neoplasm of the large intestine (ID: 12355: SCV000508204.1). A *NRAS* p.Gly12Asp (c.35G>A) 1p13.2 mutation was identified both in the primary and in the metastasis with an allele frequency of 38% and 1152 reads in the primary and an allele frequency of 30% and 130 reads in the metastasis. The mutation is listed in ClinVar database as pathogenic and is associated neoplasm of the large intestine (ID: 39648: SCV000503720.1). Two novel, not annotated *GNAS* p.Trp301* (c.902G>A) 20q13.32 and *ARID1B* p.His96del (c.270_272delCCA) 6q25.3 mutations were identified in the metastasis with an allele frequency of 5-6% with a coverage of 84 and 165 reads respectively.

4.3.1.4 Patient 10

40 years old male patient was diagnosed with CRC in the sigmoid colon with one synchronous liver metastasis in liver segment 4 (initial TNM STAGE: pT3N1M1). The patient received FOLFIRI as adjuvant treatment.

A *TP53* p.Arg175His (c.524G>A) 17p13 was identified both in the primary and in the metastasis with an allele frequency of 19% and 683 reads in the primary and an allele

frequency of 8.8% and 1973 reads in the metastasis. The mutation is listed in ClinVar database as likely pathogenic and is associated neoplasm (ID: 12374: SCV000504891.1). A *SMAD4* deletion frameshift (p.Asp415fs*20 (c.1242_1245delAGAC)) was identified in the metastasis with an allele frequency of 5% and 522 reads. The mutation is listed in COSMIC database and is associated with neoplasm in large intestine (ID: COSM1266191). A novel *TSC* splice site mutation (c.5260-1G>A 16p13.3) and a novel *GNAS* p.Trp301* (c.903G>A) 20q13.32 mutation were identified in the metastasis with an allele frequency of 5-6% with a coverage of 86 and 94 reads respectively.

4.3.1.5 Patient 11

58 years old male patient was diagnosed with CRC in the sigmoid colon with synchronous multiple liver metastases in the right lobe of the liver (initial TNM STAGE: pT3N0M1). The tumor cell percentage was 10% in the primary and 60% in the metastasis. The patient received FOLFIRI as adjuvant treatment.

A *TSC2* p.Ala1257Val (c.3770C>T) 16p13.3 mutation was identified in the primary and in the metastasis, with an allele frequency of 51% and 2294 reads in the primary and an allele frequency of 14% and 1026 reads in the metastasis. A *TP53* p.Gly245Ser (c.733G>A) 17p13 mutation was identified in the metastasis with an allele frequency 27% and 442 reads. The mutation is listed in ClinVar database as likely pathogenic and is associated with neoplasm of the large intestine (ID: 12365: SCV000504879.1). A *SMAD4* p.Asp537Gly (c.1610A>G) 18q21.2 mutation was identified in the metastasis with an allele frequency 26% and 106 reads. The somatic mutation is listed in ClinVar database as likely pathogenic and is associated with CRC (ID: 523175: SCV000734842.1). A *KRAS* p.Gly13Asp (c.38G>A) 12p12.1 was identified in the metastasis with an allele frequency of 19% and 625 reads. The mutation is listed in ClinVar database as pathogenic. The mutation is associated with neoplasm of the large intestine (ID: 12580: SCV000504458.1).

4.3.1.6 Patient 18

52 years old female patient was diagnosed with CRC in the ascending colon with synchronous liver metastasis is liver segment 2 (TNM STAGE: T4 NX pM1).

A *KRAS* p.Gly12Cys (c.34G>T) 12p12.1 mutation was identified in the primary with an allele frequency of 27% and 563 reads, and in the metastasis with 20% and 796 reads. The mutation is listed in ClinVar database as pathogenic and is associated with neoplasm of the large intestine (ID: 12578: SCV000504488.1). A *TSC2* p.Arg1483Thr (c.4448G>C) 16p13.3 mutation was identified both in the primary and in the metastasis, with an allele frequency of 26% and 1467 reads in the primary and 18% and 1444 reads in the metastasis. The mutation is listed in ClinVar database with uncertain significance and is associated with hereditary cancer-predisposing syndrome (ID: 65332: SCV000675502.2). An *ARID1B* in-frame deletion (p.His96del (c.270_272delCCA)) was identified in the metastasis with an allele frequency of 8.4% and 107 reads.

4.3.2 Cases with identical mutation pattern

4.3.2.1 Patient 3

70 years old male patient was diagnosed with CRC in the sigmoid colon. After 1 year, two metachronous liver metastases in liver segment 6 (initial TNM STAGE: pT3N2) were resected. The patient received FOLFOX, FOLFIRI and Bevacizumab as adjuvant therapy.

A *KRAS* p.Gly12Asp (c.35G>A) 12p12.1 mutation was identified both in the primary and in the metastasis with an allele frequency of 29% and 502 reads in the primary and an allele frequency of 9.6% and 741 reads in the metastasis. The mutation is listed in ClinVar database as pathogenic. The somatic mutation is associated with neoplasm of the large intestine (ID: 12582: SCV000504481.1).

4.3.2.2 Patient 19

85 years old male patient was diagnosed with CRC in the ascending colon with synchronous liver metastasis in liver segment 1 (TNM STAGE: T3G2pN2a pM1).

A *KRAS* p.GLy12Val (c.35G>T) 12p12.1 mutation was identified in the primary with an allele frequency of 54% and 631 reads and in the metastasis with an allele frequency of 43% and 862 reads (Figure 16 A). The mutation is listed in ClinVar database as pathogenic. The somatic mutation is associated with neoplasm of the large intestine (ID: 12583: SCV000203917.2). A *BRCA2* splice acceptor variant was identified both in the primary and in the metastasis (c.6842-2A>G) 13q13.1 with coverage of 53% and 58 unique reads in the

primary and a coverage of 50% and 122 unique reads in the metastasis (Figure 16 B). The mutation is listed in ClinVar database as pathogenic (ID: 491309: SCV000839907.1). A *PIK3CA* p.Asp350Gly (c.1049A>G) 3q26.32 mutation was identified in the primary with an allele frequency of 28% and 369 reads and in the metastasis with an allele frequency of 21% and 488 reads in the metastasis (Figure 16 C). The mutation is listed in ClinVar database as with uncertain significance and is associated with inborn genetic diseases (ID: 521363: SCV000741908.1).

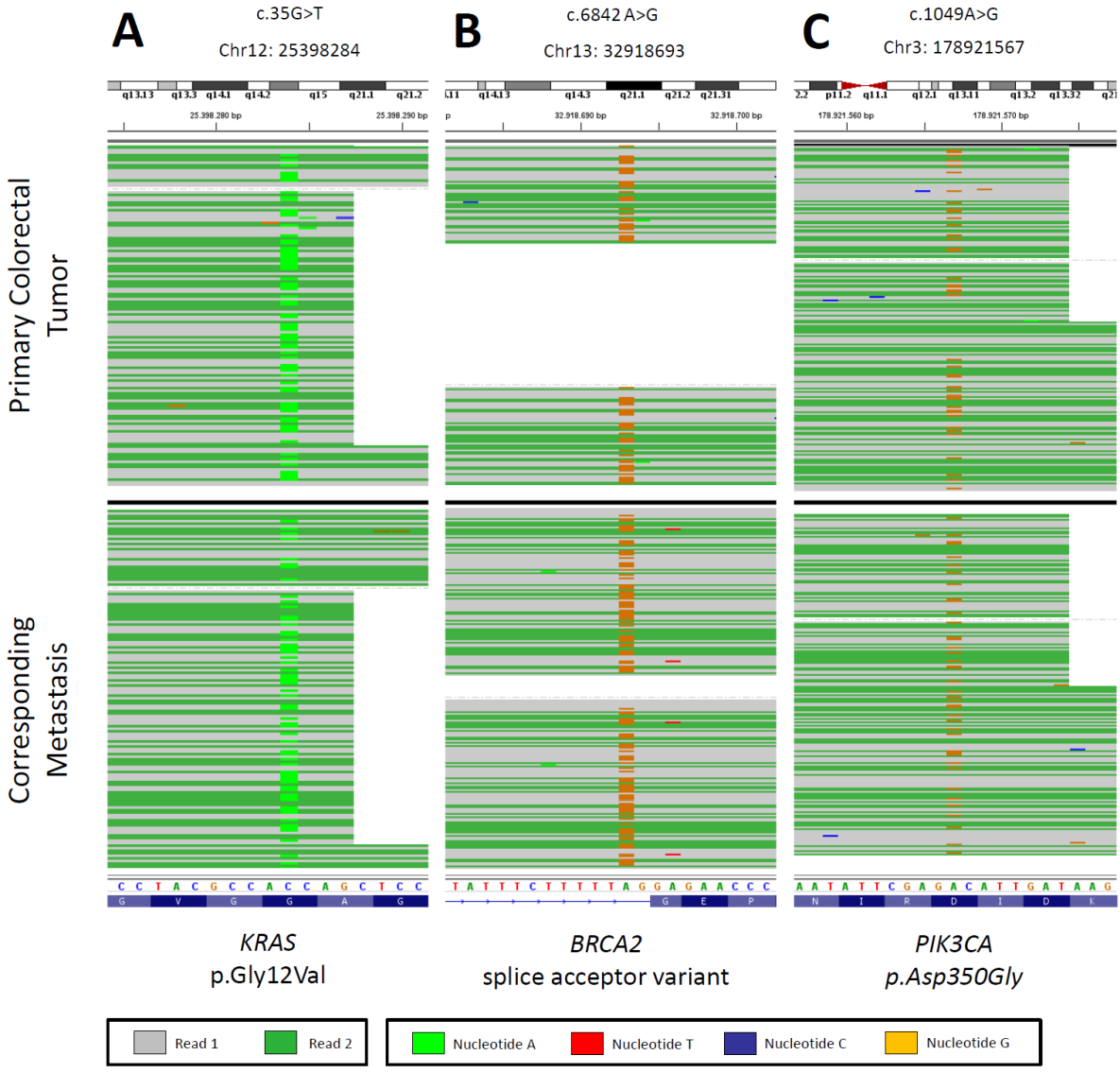

Figure 16: Exemplary raw NGS data from primary colorectal tumor (upper part) and corresponding liver metastasis (lower part) showing 3 different mutation sites. Grey or dark green lines show the two reads (forward or reverse strand). Mutations are highlighted as bars (light green for nucleotide A; red T; blue C; orange G). The single-letter amino acid code is in 5'→3'. The codons (triplets) are in A) *KRAS* 3'→5'; B) *BRCA*2 5'→3'; *PIK3CA* 5'→3' direction. A) The *KRAS* mutation (c.35G>T), B) *BRCA2* mutation (c.6824A>T) and the C) *PIK3CA* mutation (c.1049A>G) can be observed in the in the primary tumor and in the metastasis. As reference genome, GRCh37 was used.

4.3.2.3 Patient 21

69 years old male patient was diagnosed with CRC in the transverse colon. After two years, a metachronous liver metastasis in liver segment 4 (initial TNM STAGE: pT3pN0°pM0) was operated and removed. The patient received FOLFIRI FOLFOX and Bevacizumab as adjuvant treatment.

A *KRAS* p.Gly12Cys (c.34G>T) 12p12.1 mutation was identified in the primary with an allele frequency of 11% and 367 reads and in the metastasis with an allele frequency of 45% and 916 reads. The mutation is listed in ClinVar database as pathogenic and is associated with neoplasm of the large intestine (ID: 12578: SCV000504488.1). An *ARID1B* p.Asp1741del (c.5218_5220del3) 6q25.3 mutation was identified both in the primary and in the metastasis with an allele frequency 47% and 4603 reads and an allele frequency of 23% and 6455 reads respectively. The mutation is listed in the single nucleotide polymorphism database (dbSNP) with no clinical significance (ID: rs113820273).

4.3.2.4 Patient 22

51 years old male was diagnosed with CRC in the rectum. After one year, a multiplex metachronous liver metastasis in left lobe (initial TNM STAGE: pT3pN2) was surgically resected. The patient received FOLFIRI and FOLFOX as adjuvant treatment.

A *TP53* p.Leu194Pro (c.581T>C) 17p13.1 mutation was identified both in the primary and in the metastasis with an allele frequency of 54% and 2858 reads in the primary and 49% and 3307 reads in the metastasis. The mutation is listed in ClinVar database as likely pathogenic and is associated with neoplasm of the large intestine (ID: 376635: SCV000508875.1)

4.3.2.5 Patient 23

80 years old male patient was diagnosed with CRC in the sigmoid colon. After one year, a multiplex metachronous liver metastasis in liver segments 6 and 7 (initial TNM STAGE: pT4bN1c) was resected. The patient refused chemotherapy.

A *TP53* p.Ser241Phe (c.722C>T) 17p13 mutation was identified in the primary and in the metastasis, with an allele frequency of 56% and 266 reads in the primary and 44% and 316 reads in the metastasis. The mutation is listed in ClinVar database as likely pathogenic and is associated with neoplasm of the large intestine (ID: 12359: SCV000509687.1).

4.3.2.6 Patient 24

65 years old male patient was diagnosed with CRC in the sigmoid colon with two synchronous liver metastases in liver segment 4 (initial TNM STAGE: pT3N1). The patient received FOLFIRI, FOLFOX, Bevacizumab and anti-EGFR as adjuvant treatment.

A *TP53* p.Arg273Cys (c.817C>T) 17p13 mutation was identified both in the primary and in the metastasis with an allele frequency of 46% and 3094 reads in the primary and 61% and 3067 reads in the metastasis. The mutation is listed in ClinVar database as likely pathogenic and is associated with neoplasm (ID: 43594: SCV000504689.1).

4.3.2.7 Patient 16

51 years old female patient was diagnosed with CRC in the sigmoid colon with synchronous liver metastasis in liver segment 7 (initial TNM STAGE: T3G2pN2a°pM1).

A *TP53* p.Pro278Ser (c.832C>T) 17p13.1 mutation was identified both in the primary and in the metastasis with an allele frequency of 66% and 1158 reads in the primary and 30% and 3451 reads in the metastasis. The mutation is listed in ClinVar database as likely pathogenic and is associated with neoplasm of the large intestine (ID: 376642: SCV000509038.1).

4.3.2.8 Patient 28

53 years old female patient was diagnosed with CRC in the sigmoid colon. After one year, one solid metachronous liver metastasis in liver segment 7 was operated and removed (initial TNM STAGE: pT3N2b).

A *PIK3CA* p.Glu542Lys (c.1624G>A) 3q26.3 mutation was identified both in the primary and in the metastasis with an allele frequency of 3.7% and 1173 reads in the primary and with an allele frequency of 15% and 633 reads in the metastasis. The mutation is listed in ClinVar database as pathogenic and is associated with neoplasm of the large intestine (ID: 31944: SCV000503918.1).

4.3.3 Cases showing a more likely linear progression

4.3.3.1 Patient 1

65 years old female patient was diagnosed with CRC in the sigmoid colon with synchronous liver metastasis in liver segment 2 and 3 (initial TNM STAGE: T3N1cM1a). The patient received 14 cycles of Cetuximab–FOLFORI as neoadjuvant treatment.

A *TP53* p.Asp350Gly (c.1049A>G) 3q26.32 mutation was identified in the primary and the metastasis with an allele frequency of 27 % and 1177 reads in the primary and an allele frequency of 68% and 1666 unique reads in the metastasis. The mutation is listed in ClinVar database as pathogenic. The somatic mutation is associated with neoplasm of the large intestine (ID: 12364: SCV000509589.1). A *TSC1* p.Ser1043del (c.3127_3129delAGC) 9q34.1 mutation was identified in the metastasis with an allele frequency 6.9% and 678 reads. The mutation is listed in ClinVar database as benign. (ID: 64754: SCV000763724.1). Novel, not annotated *NF1* p.Leu1339_1341del (c.4015_4023del9) 17q11.2, a *PALB2* p.Cys1127fs (c.3377_3358insT) 16p12.2 and an *ARID1B* p.His96del (c.270_272delCCA) 6q25.3 mutations were identified in the primary with an allele frequency of 7-8% and a coverage ranging from 182-440.

4.3.3.2 Patient 2

77 years old male patient was diagnosed with CRC in the splenic flexure with two synchronous liver metastases within segment 3 (initial TNM STAGE: pT3N0). The patient received adjuvant treatment with FOLFIRI and FOLFOX.

A *TP53* p.Pro152fs (c.454_466del13) 17p13.1 deletion-frameshift was identified in the primary as well as in the metastasis with an allele frequency of 38% and 2917 reads in the primary and an allele frequency of 48% and 4451 reads in the metastasis. The mutation is listed in COSMIC database and is associated with neoplasm of the large intestine (ID: COSM43792). A TSC1 p.Ser1043del (c.3127_3129delAGC) 9q34.1 in frame deletion was identified in the primary with an allele frequency 5.6% and 555 reads. The mutation is listed in ClinVar database as benign (ID: 64754: SCV000763724.1). An *ARID1B* p.His96del (c.270_272delCCA) 6q25.3 mutation was identified in the metastasis with an allele frequency of 6% and 281 reads.

4.3.3.3 Patient 6

63 years old male patient was diagnosed with CRC in the rectum. After one year, two metachronous liver metastases in liver segments 6 and 8 (initial TNM STAGE: ypT3N0) were surgically resected. The patient received neoadjuvant chemo-irradiation therapy and FOLFOX as adjuvant treatment.

A *MSH2* p.Ala272Val (c.815C>T) 2p21 mutation was identified in the primary with an allele frequency of 7.5% and 80 reads. The mutation is listed in ClinVar database as with uncertain significance. The mutation is associated with colorectal cancer (ID: 41651: SCV000190347.1). A *PIK3CA* p.Thr1025Ala (c.3073A>G) 3q26.3 mutation was identified in the metastasis with an allele frequency of 59% and 145 reads. The mutation is listed in ClinVar database as pathogenic. The mutation is associated with non-small cell lung cancer (ID: 45467: SCV000062351.3). A *KRAS* p.Gly12Val (c.35G>T) 12p12.1 mutation was identified in the metastasis with an allele frequency of 56% and 678 reads in the metastasis. The mutation is listed in ClinVar database as pathogenic. The somatic mutation is associated with neoplasm of the large intestine (ID: 12583: SCV000504474.1). A *TP53* p.Arg248Gln (c.743G>A) 17p13 mutation was identified in the metastasis with an allele frequency of 53%

and 557 reads. The mutation is listed in ClinVar database as pathogenic. The somatic mutation is associated with neoplasm of the large intestine (ID: 12356: SCV000504708.1).

4.3.3.4 Patient 7

64 years old male patient was diagnosed with CRC in the rectum and a synchronous liver metastasis on the border between liver segments 2 and 3 (initial TNM STAGE: pT2N0). The patient received FOLFOX as adjuvant treatment.

KRAS was initially identified as wild type. Four splice-site mutations were identified in the metastasis *MAPK3* (c.675+1G>A) 16p11.2; *TSC2* (c.976-1G>A) 16p13.3; *TSC2* (c.2639+1G>A) 16p13.3 and *AKT1* (c.567+1G>A) 14q32.33 with an allele frequency ranging from 6-8% and 249; 284; 112 and 147 reads respectively. Only the *TSC2* (c.976-1G>A) 16p13.3 splice-site mutation is listed in ClinVar database as pathogenic in tuberous sclerosis 2 (ID: 49397: SCV000765897.1). An *ARID1A* p.Asp1900Asn (c.5698G>A) 1p36.11 mutation was identified in the metastasis with an allele frequency of 10% and 208 reads. The mutation is listed in COSMIC database as pathogenic and is associated with transitional cell carcinoma (ID: COSM254543). A *STK11* p.Pro324Leu (c.971C>T) 19p13.3 mutation was identified in the metastasis with an allele frequency of 7.9% and 101 reads. The mutation is listed in ClinVar database as likely pathogenic. The somatic mutation is associated Peutz-Jeghers syndrome (ID: 135931: SCV000510425.1). A *PTEN* p.Ala34Val (c.101C>T) 10q23.31 mutation was identified in the metastasis with an allele frequency of 5.4% and 93 reads. The mutation is listed in ClinVar database with uncertain significance (ID: 492726: SCV000692002.1). A *TSC1* p.Ser1043del (c.3127_3129delAGC) 9q34.1 in-frame deletion was identified in the metastasis with an allele frequency of 5.1% 761 reads. The mutation is listed in ClinVar database as benign (ID: 64754: SCV000763724.1). No mutations were identified in the primary CRC.

4.3.3.5 Patient 8

59 years old female patient was diagnosed with CRC in the rectum. After three years, one metachronous liver metastasis in liver segment 5 (initial TNM STAGE: ypT2N0) was resected. The patient received chemo-irradiation as neoadjuvant therapy and FOLFIRI as adjuvant treatment.

A novel *TP53* p.Gly59fs (c.176delG) 17p13.1 deletion-frameshift was identified both in the primary and in the metastasis with an allele frequency of 24% and 1313 reads in the primary and an allele frequency of 52% and 2200 reads in the metastasis. A *KRAS* p.Gly13Asp (c.38G>A) 12p12.1 was identified in the primary as well as in the metastasis with an allele frequency of 15% and 421 reads in the primary and an allele frequency of 34% and 694 reads in the metastasis. The mutation is described above (ID: 12580: SCV000504458.1). A *TSC2* (c.3815-15G>A) 16p13.3 splice site mutation was identified in the primary with an allele frequency of 7.4% and 121 reads. The mutation is listed in the ClinVar database as likely benign and is associated with tuberous sclerosis syndrome (ID: 50023: SCV000395634.2). An *EGFR* p.Asp800Asn (c.2398G>A) 7p11.2 mutation was identified in the primary with an allele frequency of 6.5% and 92 reads. The mutation is listed in the COSMIC database as pathogenic and is associated with lymphoepithelioma like carcinoma (COSM53289).

4.3.3.6 Patient 14

69 years old male patient was diagnosed with CRC in the rectum with synchronous liver metastasis in liver segment 5 (initial TNM STAGE: pT3N2). The patient did not receive neoadjuvant chemotherapy. The gland formation was 90% in the primary and 50% in the metastasis while necrosis was absent.

A *TP53* p.Gly262Val (c.785G>T) 17p13.1 mutation was identified both in the primary and in the metastasis with an allele frequency of 52% and 734 reads in the primary and an allele frequency of 44% and 100 reads in the metastasis. The mutation is listed in ClinVar database with uncertain significance and is associated with hereditary cancer-predisposing syndrome (ID: 428889: SCV000581135.3). A *KRAS* p.Gly13Asp (c.38G>A) 12p12.1 mutation was identified in the primary as well as in the metastasis with an allele frequency of 26% and 380 reads in the primary and an allele frequency 20% and 326 reads in the metastasis. The mutation is listed in ClinVar database as pathogenic and is associated with neoplasm of the large intestine (ID: 12580: SCV000504458.1). A non-annotated *PIK3CA* p.Gln981* (c.2941C>T) 3q26.32 mutation was identified in the primary with an allele frequency of 6% and 84 reads. A *TSC1* p.Ser1043del (c.3127_3129delAGC) 9q34.1 mutation was identified in the primary with an allele frequency 5.5% and 366 reads. The above described mutation is listed in ClinVar database as benign (ID: 64754: SCV000763724.1).

4.3.3.7 Patient 15

85 years old female patient was diagnosed with CRC in the transverse colon with synchronous liver metastasis in liver segment 1 (TNM STAGE: T3G2pN0 pM1).

A novel not annotated *AKT2* p.Leu363fs (c.1086_1087ins4) insertion frameshift mutation was identified as unique in the primary with an allele frequency of 11% and 54 reads.

4.3.3.8 Patient 25

67 years old male patient was diagnosed with CRC in the sigmoid colon. After one year, two metachronous liver metastases in liver segments 1 and 4 (initial TNM STAGE: pT3N2a) were surgically resected. The patient received FOLFIRI, FOLFOX and Bevacizumab as adjuvant treatment.

A *PBRM1* p.Glu1182fs (c.3543_3544insA) 3p21.1 mutation was identified both in the primary and in the metastasis with an allele frequency of 7.2% and a rather high amount of reads (7474 reads) in the primary and an allele frequency of 62% and 4768 reads in the metastasis. The mutation is listed in COSMIC database and is associated with neoplasms of the large intestine (ID: COSG93482). A *PIK3CA* p.Glu545Lys (c.1633G>A) 3q26.3 mutation was identified in the metastasis with an allele frequency of 62% and 4352 reads. The mutation is listed in ClinVar database as pathogenic and is associated with large intestine neoplasm (ID: 13655: SCV000503936.1). A *TP53* p.Gly262Val (c.785G>T) 17p13.1 mutation was identified in the primary as well as in the metastasis with an allele frequency of 6% and 818 reads in the primary and an allele frequency of 68% and 300 reads in the metastasis. The mutation is listed in ClinVar database as with uncertain significance and is associated with hereditary cancer-predisposing syndrome (ID: 428889: SCV000581135.3). A novel *ARID1B* p.His96del (c.270_272delCCA) 6q25.3 mutation was identified as unique in primary tumor with an allele frequency of 5.2% and 424 reads.

4.3.3.9 Patient 26

60 years old male patient was diagnosed with CRC in the rectum. After two years, a metachronous liver metastasis in liver segment 2 (initial TNM STAGE: pT3pN1) was resected. The patient received folinic acid (Leucovorin) and 5-Fluorouracil (5-FU) as

neoadjuvant treatment. The patient additionally received FOLFOX and Bevacizumab as adjuvant treatment.

An *ARID1B* p.Arg1552Lys (c.4655G>A) 6q25.3 mutation was identified in the primary with an allele frequency of 6.2% and 374 reads. The mutation is listed in COSMIC database as with no clinical significance (ID: COSG95505: ARID1B_ENST00000346085). A *TSC2* p.Ser1276Phe (c.3827C>T) 16p13.3 mutation was identified in the primary with an allele frequency of 9.4% and 310 reads. The mutation is listed in ClinVar database as benign (ID: 41735: SCV000675466.2).

A *KRAS* p.Gly12Asp (c.35G>A) 12p12.1 mutation was identified in the metastasis with an allele frequency of 38% and 3379 reads. The mutation is listed in ClinVar database as pathogenic. The somatic mutation is associated with neoplasm of the large intestine (ID: 12582: SCV000504481.1). A MLH1 p.Arg265Cys (c.793C>T) 3p22 mutation was identified in the metastasis with an allele frequency of 8.9% and 282 reads. The mutation is listed in ClinVar database as pathogenic and is associated with hereditary cancer-predisposing syndrome (ID: 29654: SCV000276047.4).

4.3.3.10 Patient 27

56 years old male patient was diagnosed with CRC in the rectum. After two years, a multiplex metachronous liver metastasis in the left lobe was operated and removed (initial TNM STAGE: pT3pN0). The patient received FOLFIRI and Bevacizumab as adjuvant treatment

A *TP53* p.Arg248Trp (c.742C>T) 17p13 mutation was identified in the primary as well as in the metastasis with an allele frequency of 27% and 350 reads in the primary and with an allele frequency of 29% and 782 reads in the metastasis. The mutation is listed in ClinVar database as likely pathogenic and is associated with neoplasm of the large intestine (ID: 12347: SCV000504850.1). A novel *ARID1B* p.His96del (c.270_272delCCA) 6q25.3 mutation was identified as unique in primary tumor with an allele frequency of 5.9% and 471 reads.

4.3.3.11 Patient 20

A 52 years old male patient was diagnosed with CRC in the sigmoid colon with synchronous liver metastasis (initial TNM STAGE: pT3N2aM1a). The patient received FOLFIRI, FOLFOX and Bevalizumab as adjuvant treatment.

A *TP53* p.Arg196Ter (c.586C>T) 17p13 mutation was identified both in the primary and in the metastasis with an allele frequency of 24% and 2961 reads in the primary and an allele frequency of 64% and 1101 reads in the metastasis. The mutation is listed in ClinVar database as pathogenic and is associated with hereditary cancer-predisposing syndrome (ID: 43589: SCV000186503.6). A *KRAS* p.Gln22Lys (c.64C>A) 12p12.1 mutation was identified in the primary with an allele frequency of 22% and 789 reads. The mutation is listed in ClinVar database as likely pathogenic and is associated with neoplasm of the large intestine (ID: 376325: SCV000505665.1). No reads were identified for *KRAS* 12q12.1 gene in the metastasis.

4.3.3.12 Patient 30

64 years old male patient was diagnosed with CRC in the ascending colon with synchronous liver metastasis in liver segment 4 (TNM STAGE: T3G2pN1b(2/24) pM1).

A *TP53* p.Lys132Asn (c.396G>T) 17p13.1 mutation was identified in the primary as well as in the metastasis, with an allele frequency of 52% and 704 reads in the primary and an allele frequency of 69% and 628 reads in the metastasis. The mutation is listed in ClinVar database as likely pathogenic and is associated with neoplasm of the large intestine (ID: 376624: SCV000508730.1). A homozygote *BRCA2* p.Asn2135fs (c.6402_6406del5) 13q13 mutation was identified in the primary with an allele frequency of 88% and 273 reads. No reads of the *BRCA* 13q13 gene were detected in the metastasis.

5. Discussion

There are mainly two scientifically recognized/accepted models of tumor metastasis, the linear progression model according to *Fearon* and *Vogelstein* (Fearon & Vogelstein, 1990) and the parallel progression model according to *Klein* (Klein, 2009) and *Hamilton* (Hamilton & Rath, 2018) (Figure 3). The linear model describes the progression of healthy tissue to adenoma and ultimately (Figure 2) to metastatic disease via dysplasia. In contrast, the parallel model postulates that metastatic founder cells mostly disseminate before the disease is clinically detected. In this project some of the cases did not fulfill the inclusion criteria neither for the parallel nor for the linear model; therefore, they can be considered debatable and are referred to as cases showing a more likely linear progression.

Tumors are characterized by tremendous heterogeneity regarding their morphological, phenotypic and genetic profiles. Therefore, when analyzing genetic aberrations, a highly sensitive detection method should be used in order to detect and identify genetic mutations that could be of great clinical interest and significance (Blank et al., 2018; Noorbakhsh et al., 2018). In the present study, targeted NGS via Illumina was used for mutation detection with a test panel covering 47 genes that are implicated in cancer (see Material and Methods part). NGS enables the screening of multiple mutations in multiple genes simultaneously; furthermore, it has high sensitivity and speed. Reproducibility and repeatability are also a major advantage of NGS, necessary for consistent allele detection (D'Haene et al., 2018; D'Haene et al., 2015). However, in the case of two patients (patients 20 and 30), a non-negligible fraction of reads of *KRAS* and *BRCA2* were left unmapped (missing reads). Missing reads can lead to inaccurate tumor progression characterization. Sample contamination or even a sequencing error could justify the missing reads (Gulilat et al., 2019; Hasan et al., 2019). Further, the test panel capacity could constitute a technical limitation regarding the characterization of the tumor progression since important genes related to CRC such as *APC* were not included. In addition to the technical challenges mentioned above, in a few cases the amount of detected reads was highly discrepant to the amount of the respective complementary reads; this could be explained by a low binding efficiency of the polymerase (Figure 15).

Tumors often have genetic and epigenetic alterations but it is not yet fully understood which of those alterations cause a survival advantage and which not. Such genetic alterations are driver and passenger mutations. Driver mutations induce cell proliferation and tumor growth, while passenger mutations have no effect on tumor development. Passenger mutations represent approximately 97% of all cancerous mutations and are important indicators regarding the cancer evolution timeline in individual patients (Vogelstein et al., 2013; Yachida et al., 2010). In patients 1, 2, 4-8, 10, 12, 13, 18, 25-27 and 29 following mutations are considered passenger mutations either due to the low allele frequency or due to the lack of clinical relevance: *ARID1B* c.270_272delCCA; *AKT1:* c.567+1G>A; *BRCA2:* c.3854delA; *EGFR:* c.2398G>A; *MAPK3:* c.675+1G>A; *MSH2:* c.815C>T; *NF1:* c.1658A>G, c.4015_4023del9, c.1400C>T; *PALB2*: c.3377_3358insT; *PTEN:* c.101C>T; *SMAD4:* c.1242_1245del4, c.1255G>T; *STK11:* c.971C>T; *GNAS:* c.903G>A; *TSC1:* c.3127_3129delAGC and *TSC2:* c.5260-1G>A, c.3770C>T, c.976-1G>A, c.5260-1G>A, c.2639+1G>A, c.3815-1G>A, c.3827C>T. It was shown that depletion of *ARID1B* in *ARID1A*-proficient CRC cell lines reduces proliferation and causes destabilization of the chromatin remodeling complex SWI/SNF (Niedermaier et al., 2019; Savas & Skardasi, 2018), a fact that could be interpreted as a possibility of the *ARID1B* mutation being a driver mutation. Mutations with an allele frequency <2% can be difficult to be detected. In addition, low allele frequency could be an indication that the specific mutation is present only in a minimal amount of cells rather than in the whole tumor entity. Mutations with allele frequency >10% are of clinical significance, therefore only these mutations were taken into consideration when evaluating the data for the progression models.

The pattern of genomic mutations that were identified in this project (cohort A) and are listed in Table 8 is mostly consistent with prior studies of CRC (Ahlquist et al., 2008; AlDubayan et al., 2018; Baba et al., 2010; Baran et al., 2018; Barber et al., 2004; Cajuso et al., 2014; De Roock et al., 2011; Donehower et al., 2013; N. I. Fleming et al., 2013; Francipane & Lagasse, 2014; Fumagalli et al., 2010; Hamada et al., 2017; Irahara et al., 2010; Y. S. Kim et al., 2017; Lin et al., 2015; Malapelle et al., 2016; Oh et al., 2018; Rychahou et al., 2008; Wilson et al., 2010). The mutation rate of genes which are of great significance in CRC such as *TP53, KRAS, PIK3CA, PTEN* and *SMAD4* was consistent with the mutation rates in the literature (Table 8) indicating that cohort A is a representative cohort for the investigation of genetic

alterations in CRC. Interestingly, in another type of cancer, non-small cell lung cancer *TP53, KRAS and PIK3CA* are also found mutated but have lower mutation rates than in CRC, namely 44.9%, 29.6 and 9% respectively (Zhao et al., 2019). The mutation rate of other genes from the present study such as *NRAS* and *TSC2* was rather inconsistent in comparison to the literature. *NRAS* had in this study a mutation rate of 3% whereas the mutation rate found in the literature is 15%. The mutation rate of *TSC2* in the current study was 10% which is the tenfold (<1%) of prior studies. These discrepancies could be explained due to the small size of cohort A in comparison to the cohorts that were used for meta-analysis by other research groups (Francipane & Lagasse, 2014; Irahara et al., 2010; Zhunussova et al., 2019).

Analyzing cohort A primary CRC, unique mutations were identified in the following genes: *ARID1B; SMAD4; BRCA2; NF1; TSC1; TSC2; AKT2; MAPK3; MSH2; MLH1; PALB2; ARID1A; PIK3CA; KRAS* and *TP53.* These genes are mainly linked to four different hallmarks of cancer: sustaining proliferative signaling (proliferation); evading apoptosis; insensitivity to anti-growth signals and genomic instability (Hanahan & Weinberg, 2011) (Figure 17). Their exact role in CRC is described in Table 8. Brannon AR *et al.* (Brannon et al., 2014) who performed a similar study to the present one and sequenced primary CRC as well as the corresponding liver metastasis, found unique mutations in the primary tumor amongst others in the following two genes: *PIK3CA* and *TP53.* These two genes were also mutated in the primary tumor in this study. Only in one patient no mutations were detected neither in the metastasis nor in the primary. To date, there are not many studies available that investigated both CRC tumors and liver metastasis in the same patient; therefore, the aforementioned mutations from the present study are a rather new finding.

Some of the gene mutations identified in the cohort A tumors were not annotated in neither of the databases that were used for interpretation. The following non annotated mutations were identified: *GNAS* (c.902G>A); *ARID1B* (c.270_272delCCA); *NF1* (c.4015_4023del9); *PALB2* (c.3377_3358insT); *PIK3CA* (c.2941C>T); *TSC1* (c.3127_3129delAGC) and *AKT2* (c.1086_1087ins4). Lack of annotation does not necessarily mean that the mutation has no clinical significance. GNAS protein plays an important role in cell proliferation and thus in tumor development (S. M. Bray et al., 2019). *GNAS* p.Trp301* (c.902G>A) mutation introduces an immediate translation stop codon that causes premature termination of the protein that may lead to change or total loss of function of the protein. This mutation could

be proven to be clinically beneficial for stalling tumor growth. The *ARID1B* c.270_272delCCA is a deletion of histidine. It is rather difficult to assess whether this deletion could have a clinical effect in CRC without experimental data. *NF1* is a negative regulator of *KRAS* (Ahlquist et al., 2008). *NF1* (c.4015_4023del9) mutation and *TSC1* (c.3127_3129delAGC) mutation are deletions for which is unknown whether they affect the clinical outcome in CRC patients. *PALB2* c.3377_3358insT and *AKT2* c.1086_1087ins4 are frameshift mutations. In general, this type of mutations has severe effects and considering the fact that *PALB2* is involved in DNA damage repair (AlDubayan et al., 2018) and *AKT2* in proliferation (Agarwal et al., 2017), it can be hypothesized that the aforementioned mutations are of clinical relevance. *PIK3CA* c.2941C>T leads to loss or change of protein function and since *PIK3CA* is an oncogene such a tremendous functional deterioration could have an effect in tumor development (Samuels et al., 2004).

In cohort A, it was investigated whether the histotype, resection time, localization or necrosis status were associated with certain gene mutations. A significant association was only observed for *MLH1* with the mucinous histotype and for *PTEN* and *TP53* mutations with the non-mucinous adenocarcinoma (p=0.0142; p=0.0046; p=0.036 respectively). These results are in concordance with the literature. *MLH1* was shown to be frequently mutated in MSI tumors; these tumors are associated with mucinous histotype (Armaghany et al., 2012). *Reynolds et al.* revealed that mutated *TP53* is found more frequently in the non-mucinous histotype (Reynolds et al., 2019). In line with *Madsen et al*. who demonstrated in human induced pluripotent stem cells that mutated *PIK3CA* tumors exhibit extensive necrosis, the present study also showed that *PIK3CA* mutation was significantly associated with the necrosis level (Madsen et al., 2019). In addition, it was investigated whether certain gene mutations are more prominent in either the linear or the parallel model. *SMAD4*, *ARID1A* and *PTEN* mutations were significantly more in tumors of the parallel model, indicating an association between the mutated genes and the model of progression. This finding is rather new since this association was not reported before.

In cohort B, four gene mutations (*KDM6A, MSH6, STK11, TSC1*) in the primary tumor and two gene mutations (*MSH6, TSC1)* in the metastatic tumor were only detected once and were significantly associated (Fisher test: p= 0.00076) with the progression models. All 59 cases referring to linear progression were classified correctly. From 7 cases referring to parallel

progression 57% were classified correctly. Although this association was statistically significant, it cannot be considered plausible since only one mutation is not sufficient for drawing an accurate conclusion. In both cohorts, it was observed that at an allele frequency >10% a high mutation rate was an indication for parallel progression.

In four cases of cohort A (patients 5, 9, 17 and 29), the same mutations were found in the primary tumor as well as in the metastasis indicating that both tumors originate from the same precursor clone. Such mutations were found in the following genes: *PTEN, ARID1A, TP53, KRAS* (Table 8). In two of these cases, additional mutations were detected only in the liver metastasis in the following genes *TP53, KRAS, PIK3CA* which underlies their role in tumor progression (Lipsyc & Yaeger, 2015) (Table 8). Furthermore, in these four cases certain mutations in *SMAD4, ARID1A* and *TP53* were detected only in the primary tumor indicating tumor dissemination at an early stage which refers to the parallel progression model (Table 8). Altogether, the genetic landscape of the tumors of these patients corresponds clearly to the parallel metastasis progression model. In addition, regarding cohort B, 14 out of 69 patients had unique mutations in the primary tumor, which is a similar ratio to the ratio of cohort A (4 out of 30). *Kim et al*. (R. Kim et al., 2015) investigated the genetic mutations in the primary CRC tumor and in the corresponding liver metastases of eight patients. Three of them had mutations in the primary tumor that could not be found in the metastasis. The ratio of patients with unique primary tumor mutations in cohort A (4 out of 30) is lower than the ratio that *Kim et al.* report (3 out of 8), probably due to the small sample size of the latter.

Fourteen cases were allocated to the linear progression model due to the presence of the same gene mutation in both the primary tumor and the metastasis. Six patients had additional mutations in the metastasis. The additional mutations lead to gain or loss of function enabling cancer invasion and spreading metastasis into other tissue (R. Kim et al., 2015; Suzuki & Tarin, 2007). In seven patients, discrepant mutations were not identified in the following five genes, *KRAS, TP53, BRCA2, PIK3CA* and *ARID1B*, neither in the primary tumor nor in the metastasis. Excluding from the data evaluation all the mutations with an allele frequency lower than 5% leads to exclusion of most of the gene mutations present in the tumor (Bozic et al., 2016). This means that the two tumors (primary and metastasis) are genetically similar. In this case, a metastasis can occur via an epithelial to mesenchymal transition triggered by

changes in gene expression that do not necessarily result from mutations but rather from physiological and environmental alterations (Chaffer & Weinberg, 2011; R. Kim et al., 2015). Of note, patient 19 had *KRAS, PIK3CA and BRCA2* mutations in the primary tumor and in the metastasis (Figure 16). *BRCA2* mutations lead to DNA damages (Mairinger et al., 2017) which can cause the additional driver gene mutations in *KRAS* and *PIK3CA* that were probably acquired at the invasive phase enabling the metastasis (Vogelstein & Kinzler, 2015).

For ten cases (cohort A) is not clear whether they show linear or parallel progression and are therefore mentioned as cases showing a more likely linear progression. Although unique mutations were identified in the primary tumor (described and discussed above), these cases probably correspond to the linear progression model since they have a low allele frequency (<10%) and/or no clinical significance. In other words, primary tumor and metastasis are genetically similar but not identical since for the primary tumor to metastasize, mutagenesis leading to gain or loss of function is necessary (R. Kim et al., 2015; Klein, 2009; Montel et al., 2006). In two cases (patients 20 and 30), although the allele frequency in the primary tumor was higher than 10%, no allele was identified in the metastasis, either due to technical issues as described above or gene loss. Interestingly, in one patient no common mutations between the primary tumor and the metastasis could be identified due to very low sequence coverage. In addition, the *NF1* mutation in the primary tumor of this patient had a low allele frequency. Heterogeneity of the primary tumor could explain the lack of common mutations (Alizadeh et al., 2015). Of note, this study cannot differentiate between tumor cells that remain in a quiescent state until the environmental conditions are favorable to begin proliferation again (dormancy model) and tumor cells that continue their proliferative cycle (Neophytou et al., 2019).

Table 8: List of mutated genes of our cohort compared to literature. Mutation Rate (MR) >10% refers to the allele frequency >10%. MR with MSI refers to reference cohorts including only microsatellite instability high tumors. All other reference MR refers to chromosomal instable (CIN) CRC. (Barber et al., 2004; De Roock et al., 2011; Irahara et al., 2010; Malapelle et al., 2016; Wilson et al., 2010)

Gene	Protein	Function	MR (Total)	MR >10%	MR rev.	
ARID1A	AT-rich interactive domain-containing protein 1A	Chromatin remodeling	10%	3%	5%(39%MSI)	(Baran et al., 2018; Cajuso et al., 2014)
ARID1B	AT-rich interactive domain-containing protein 1B	Chromatin remodeling	37%	3%	13%MSI	(Cajuso et al., 2014)
BRCA2	BRCA2, DNA repair associated	DNA damage repair	10%	7%	4%	(Oh et al., 2018)
MLH1	MutL homolog 1	DNA damage repair	3%	3%	75%MSI	(Donehower et al., 2013)
MSH2	DNA mismatch repair protein Msh2	DNA damage repair	3%	0%	40%MSI	(Lin et al., 2015)
NF1	Neurofibromin 1	Cell proliferation and survival	10%	0%	4%	(Ahlquist et al., 2008)
PALB2	Partner and localizer of BRCA2	DNA damage repair	3%	0%	0.5%	(AlDubayan et al., 2018)
PBRM1	Protein polybromo-1	Chromatin remodeling	3%	3%	N.A	
PTEN	Phosphatase And Tensin Homolog	Cell cycle progression and survival	7%	3%	5-14%	(De Roock et al., 2011)
SMAD4	SMAD Family Member 4	Cell proliferation	20%	10%	9%	(N. I. Fleming et al., 2013)
STK11	Serine/threonine kinase 11	Apoptosis and DNA damage	7%	0%	0.8%	(Malapelle et al., 2016)
TP53	Cellular tumor antigen p53	Cell cycle arrest, apoptosis	73%	73%	59%(20%MSI	(Baran et al., 2018; Cancer Genome Atlas, 2012)
TSC1	TSC Complex Subunit 1	Cell growth	7%	0%	1%	(Francipane & Lagasse, 2014)
TSC2	TSC Complex Subunit 2	Cell growth	10%	10%	1%	(Francipane & Lagasse, 2014)
AKT1	AKT Serine/Threonine Kinase 1	Proliferation, survival and angiogenesis	3%	0%	1-6%	(Fumagalli et al., 2010)
AKT2	AKT Serine/Threonine Kinase 2	Proliferation, survival and angiogenesis	3%	3%	N.A	
EGFR	Epidermal Growth Factor Receptor	Cell proliferation, apoptosis	3%	0%	<1%	(Barber et al., 2004)
GNAS	GNAS Complex Locus	Cell proliferation	7%	0%	9%	(Wilson et al., 2010)
KRAS	KRAS Proto-Oncogene, GTPase	Cell proliferation and survival	43%	43%	40%-44%	(Baran et al., 2018; Fumagalli et al., 2010)
MAPK3	Mitogen-Activated Protein Kinase 3	Cell growth, adhesion and survival	3%	0%	N.A	
NRAS	NRAS proto-oncogene, GTPase	Cell proliferation and survival	3%	3%	15%	(Irahara et al., 2010)
PIK3CA	PI3-Kinase P110 Subunit Alpha	Cell proliferation and survival	17%	17%	10-20%	(Hamada et al., 2017)

Table 9: Overview of the mutations with an allele frequency >10%. For mutations in [], no allele was identified in the tumor.

		Patient	Mutations with minimum 10% allele frequency		
			mutual	unique in primary	unique in metastasis
Parallel Progression	unique mutation in primary	17	PTEN	SMAD4	KRAS TP53
		9	ARID1A TP53	ARID1A	-
		5	KRAS	TP53	PIK3CA TP53
		29	KRAS	SMAD4	-
Linear Progression	additional mutations in metastasis	12	KRAS SMAD4 TP53	-	TSC2
		4	TP53	-	-
		13	NRAS TP53	-	-
		10	TP53	-	-
		11	TSC2	-	KRAS TP53 SMAD4
		18	KRAS TSC2	-	-
	identical mutation pattern	3	KRAS	-	-
		19	BRCA2 KRAS PIK3CA	-	-
		21	ARID1B KRAS	-	-
		22	TP53	-	-
		23	TP53	-	-
		24	TP53	-	-
		16	TP53	-	-
		28	PIK3CA	-	-
	more likely linear progression	1	TP53	-	-
		2	TP53	-	-
		6	-	-	KRAS PIK3CA TP53
		7	-	-	-
		8	KRAS TP53	-	-
		14	KRAS TP53	-	-
		15	-	AKT2	-
		30	TP53	BRCA2	[BRCA2]
		20	TP53	KRAS	[KRAS]
		25	PBRM1 TP53	-	PIK3CA
		26	-	-	KRAS MLH1
		27	TP53	-	-

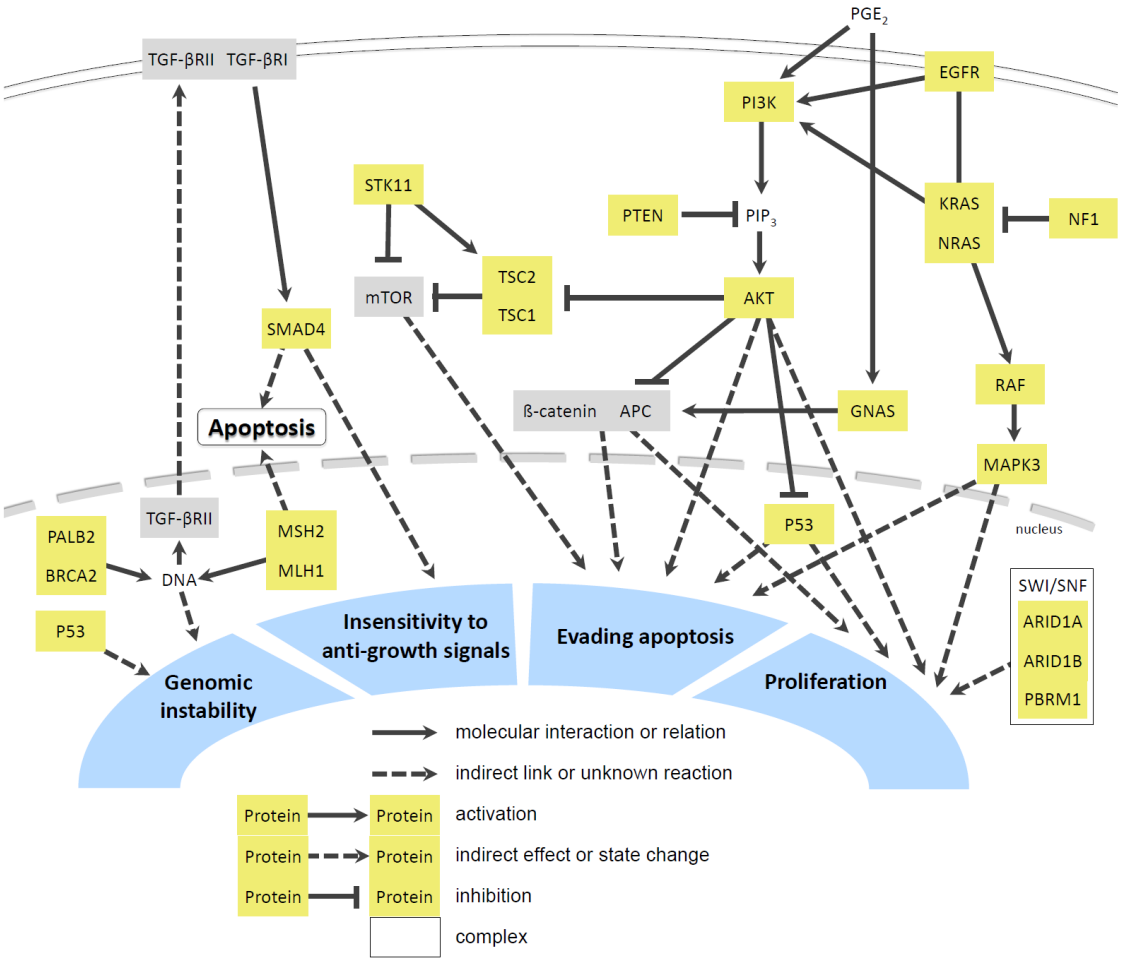

Figure 17: Simplified cancer pathway. Genes mutated in this study are marked as yellow. Adapted from Kyoto Encyclopedia of Genes and Genomes (KEGG): 05200 Pathways in cancer; 04350 TGF-beta signaling pathway; 04150 mTOR-signaling pathway; 05225 hepatocellular carcinoma pathway.

.

Abstract

Introduction: Currently two major progression schemes are proposed to explain the development of metastasis. The linear model describes the origin of distant metastasis arising from cells of primary tumor spreading after sub clonal evolution via malignant disseminated tumor cells. In contrast, the parallel model proposes an early dissemination of the same precursor clone and independent progression of the primary tumor and metastasis. In order to identify underlying mechanisms in colorectal cancer (CRC) we compared the mutational profile of corresponding primary and metastatic tumor samples.

Materials and Methods: Formalin-fixed paraffin embedded (FFPE) tissues from primary (CRC) samples (n=30) and their corresponding liver metastasis (n=30) were obtained from 2 different institutions. Isolated DNA from the aforementioned samples was subjected to targeted amplicon sequencing using the Illumina MiSeq platform. The custom-designed panel covers hotspots and whole coding sequences of 45 cancer related genes. Findings were validated using an independent cohort of 69 pairs obtained from the cancer genome atlas (TCGA).

Results: Within our cohort in primary tumors, 57 mutations were identified of which 9 were unique in the primary. Additionally, in corresponding liver metastasis 69 mutations were identified, 33 of them unique at the metastatic site. 4/30 patients had additional mutations in the primary. 8/30 patients showed an identical mutation pattern and 6/30 patients had additional mutations in the metastasis indicating a linear progression. 12/30 were debatable, indicating a more likely linear progression. Unique mutations both in primary and metastatic tumor were found in 22 patients. In the TCGA cohort, 14/69 cases were classified as parallel.

Conclusion: In our study, we demonstrated that in a significant proportion of CRC the metastasis progressed parallel with the primary tumor. The impact of different progression models should be taken into consideration for improving current molecular diagnostics and when developing personalized therapeutic strategies.

Zusammenfassung

Einleitung:

Derzeit werden zwei unterschiedliche Modelle zur Entstehung von Metastasen diskutiert. Das lineare Modell beschreibt die Fähigkeit der Tumore zur Metastasierung nach subklonaler Evolution einzelner Zellen des Primärtumors. Im Gegensatz dazu geht das parallele Modell von einer frühen Dissemination einer gemeinsamen malignen Ursprungszelle aus, welche jeweils eine unabhängige Progression zeigt. Beide Modelle können aufgrund unterschiedlicher molekularer Eigenschaften unterschieden werden und die Identifikation kann dabei helfen eine geeignete Therapieauswahl zu treffen. Ziel dieser Arbeit war es Mutationsprofile von primären kolorektalen Karzinomen (CRC) zu identifizieren und mit korrespondierenden metastasierter Tumoren zu vergleichen.

Methoden:

Aus jeweils 30-Formalin-fixierten Paraffin-eingebetteten Geweben von kolorektalen Karzinomen und den korrespondierenden Lebermetastasen wurde DNA isoliert. Zur Identifikation der jeweiligen Genotypen wurde diese dann mittels zielgerichteter Amplicon-Sequenzierung unter Verwendung eines Tumorpanels (45 tumorassoziierte Gene), analysiert. Zur Validierung wurde eine Kohorte mit 69 Patienten aus dem TCGA-Projektes verwendet.

Ergebnisse:

In den Primärtumoren wurden insgesamt 57 Mutationen identifiziert, von denen 9 ausschließlich im Primarius gefunden wurden. In den Metastasen konnten 69 Mutationen detektiert werden von denen 33 nicht im Primärtumor nachweisbar waren. Bei 8 Patienten wurde das exakt gleiche Mutationsmuster im Primärtumor und in der Metastase gefunden. Bei 21 Patienten wurden einzigartige Mutationen entweder im Primarius (4; 14/69 TCGA), was für das parallele Tumorprogressionsmodell spricht, oder in der Metastase gefunden.

Schlussfolgerung:

In der Studie konnte gezeigt werden, dass ein signifikanter Anteil der Metastasierung parallel zum Primärtumor voranschreitet. Eine Analyse des molokularen Status von sowohl Primarius als auch der Metastase könnte helfen, die derzeitige Molekulardiagnostik zu verbessern und personalisierte Therapiestrategien weiterzuentwickeln.

6. References

1. Aaltonen, L. A., Hamilton, S. R., World Health Organization.,International Agency for Research on Cancer. (2000): Pathology and genetics of tumours of the digestive system. Lyon Oxford: IARC Press ; Oxford University Press (distributor).

2. Agarwal, E., Robb, C. M., Smith, L. M., Brattain, M. G., Wang, J., Black, J. D.,Chowdhury, S. (2017): Role of Akt2 in regulation of metastasis suppressor 1 expression and colorectal cancer metastasis. Oncogene 36, 3104-3118.

3. Ahlquist, T., Bottillo, I., Danielsen, S. A., Meling, G. I., Rognum, T. O., Lind, G. E., Dallapiccola, B.,Lothe, R. A. (2008): RAS signaling in colorectal carcinomas through alteration of RAS, RAF, NF1, and/or RASSF1A. Neoplasia 10, 680-686, 682 p following 686.

4. AlDubayan, S. H., Giannakis, M., Moore, N. D., Han, G. C., Reardon, B., Hamada, T., Mu, X. J., Nishihara, R., Qian, Z., Liu, L., Yurgelun, M. B., Syngal, S., Garraway, L. A., Ogino, S., Fuchs, C. S.,Van Allen, E. M. (2018): Inherited DNA-Repair Defects in Colorectal Cancer. Am J Hum Genet 102, 401-414.

5. Alizadeh, A. A., Aranda, V., Bardelli, A., Blanpain, C., Bock, C., Borowski, C., Caldas, C., Califano, A., Doherty, M., Elsner, M., Esteller, M., Fitzgerald, R., Korbel, J. O., Lichter, P., Mason, C. E., Navin, N., Pe'er, D., Polyak, K., Roberts, C. W., Siu, L., Snyder, A., Stower, H., Swanton, C., Verhaak, R. G., Zenklusen, J. C., Zuber, J.,Zucman-Rossi, J. (2015): Toward understanding and exploiting tumor heterogeneity. Nat Med 21, 846-853.

6. Armaghany, T., Wilson, J. D., Chu, Q.,Mills, G. (2012): Genetic alterations in colorectal cancer. Gastrointest Cancer Res 5, 19-27.

7. Arnold, C. N., Goel, A., Blum, H. E.,Boland, C. R. (2005): Molecular pathogenesis of colorectal cancer: implications for molecular diagnosis. Cancer 104, 2035-2047.

8. Baba, Y., Nosho, K., Shima, K., Meyerhardt, J. A., Chan, A. T., Engelman, J. A., Cantley, L. C., Loda, M., Giovannucci, E., Fuchs, C. S.,Ogino, S. (2010): Prognostic significance of AMP-activated protein kinase expression and modifying effect of MAPK3/1 in colorectal cancer. Br J Cancer 103, 1025-1033.

9. Baptistella, A. R., Salles Dias, M. V., Aguiar, S., Jr., Begnami, M. D.,Martins, V. R. (2016): Heterogeneous expression of A33 in colorectal cancer: possible explanation for A33 antibody treatment failure. Anticancer Drugs 27, 734-737.

10. Baran, B., Mert Ozupek, N., Yerli Tetik, N., Acar, E., Bekcioglu, O.,Baskin, Y. (2018): Difference Between Left-Sided and Right-Sided Colorectal Cancer: A Focused Review of Literature. Gastroenterology Res 11, 264-273.

11. Barber, T. D., Vogelstein, B., Kinzler, K. W.,Velculescu, V. E. (2004): Somatic mutations of EGFR in colorectal cancers and glioblastomas. N Engl J Med 351, 2883.

12. Blank, A., Roberts, D. E., 2nd, Dawson, H., Zlobec, I.,Lugli, A. (2018): Tumor Heterogeneity in Primary Colorectal Cancer and Corresponding Metastases. Does the Apple Fall Far From the Tree? Front Med (Lausanne) 5, 234.

13. Blausen.com. (2014): Medical gallery of Blausen Medical: WikiJournal of Medicine 1;ISSN:2002-4436;DOI:10.15347/wjm/2014.010.

14. Bozic, I., Gerold, J. M.,Nowak, M. A. (2016): Quantifying Clonal and Subclonal Passenger Mutations in Cancer Evolution. PLoS Comput Biol 12, e1004731.

15. Brannon, A. R., Vakiani, E., Sylvester, B. E., Scott, S. N., McDermott, G., Shah, R. H., Kania, K., Viale, A., Oschwald, D. M., Vacic, V., Emde, A. K., Cercek, A., Yaeger, R., Kemeny, N. E., Saltz, L. B., Shia, J., D'Angelica, M. I., Weiser, M. R., Solit, D. B.,Berger, M. F. (2014): Comparative sequencing analysis reveals high genomic concordance between matched primary and metastatic colorectal cancer lesions. Genome Biol 15, 454.

16. Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A.,Jemal, A. (2018): Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 68, 394-424.

17. Bray, S. M., Lee, J., Kim, S. T., Hur, J. Y., Ebert, P. J., Calley, J. N., Wulur, I. H., Gopalappa, T., Wong, S. S., Qian, H. R., Ting, J. C., Liu, J., Willard, M. D., Novosiadly, R. D., Park, Y. S., Park, J. O., Lim, H. Y., Kang, W. K., Aggarwal, A., Kim, H. C.,Reinhard, C. (2019): Genomic characterization of intrinsic and acquired resistance to cetuximab in colorectal cancer patients. Sci Rep 9, 15365.

18. Brierley, J., Gospodarowicz, M. K.,Wittekind, C. TNM classification of malignant tumours (Eighth edition).

19. Cajuso, T., Hanninen, U. A., Kondelin, J., Gylfe, A. E., Tanskanen, T., Katainen, R., Pitkanen, E., Ristolainen, H., Kaasinen, E., Taipale, M., Taipale, J., Bohm, J., Renkonen-Sinisalo, L., Mecklin, J. P., Jarvinen, H., Tuupanen, S., Kilpivaara, O.,Vahteristo, P. (2014): Exome sequencing reveals frequent inactivating mutations in ARID1A, ARID1B, ARID2 and ARID4A in microsatellite unstable colorectal cancer. Int J Cancer 135, 611-623.

20. Cancer Genome Atlas, N. (2012): Comprehensive molecular characterization of human colon and rectal cancer. Nature 487, 330-337.

21. Carrato, A. (2008): Adjuvant treatment of colorectal cancer. Gastrointest Cancer Res 2, S42-46.

22. Caswell, D. R.,Swanton, C. (2017): The role of tumour heterogeneity and clonal cooperativity in metastasis, immune evasion and clinical outcome. BMC Med 15, 133.

23. Chaffer, C. L.,Weinberg, R. A. (2011): A perspective on cancer cell metastasis. Science 331, 1559-1564.

24. Chakravarty, D., Gao, J., Phillips, S. M., Kundra, R., Zhang, H., Wang, J., Rudolph, J. E., Yaeger, R., Soumerai, T., Nissan, M. H., Chang, M. T., Chandarlapaty, S., Traina, T. A., Paik, P. K., Ho, A. L., Hantash, F. M., Grupe, A., Baxi, S. S., Callahan, M. K., Snyder, A., Chi, P., Danila, D., Gounder, M., Harding, J. J., Hellmann, M. D., Iyer, G., Janjigian, Y., Kaley, T., Levine, D. A., Lowery, M., Omuro, A., Postow, M. A., Rathkopf, D., Shoushtari, A. N., Shukla, N., Voss, M., Paraiso, E., Zehir, A., Berger, M. F., Taylor, B. S., Saltz, L. B., Riely, G. J., Ladanyi, M., Hyman, D. M., Baselga, J., Sabbatini, P., Solit, D. B.,Schultz, N. (2017): OncoKB: A Precision Oncology Knowledge Base. JCO Precis Oncol 2017

25. Cserepes, M., Ostoros, G., Lohinai, Z., Raso, E., Barbai, T., Timar, J., Rozsas, A., Moldvay, J., Kovalszky, I., Fabian, K., Gyulai, M., Ghanim, B., Laszlo, V., Klikovits, T., Hoda, M. A., Grusch, M., Berger, W., Klepetko, W., Hegedus, B.,Dome, B. (2014): Subtype-specific KRAS mutations in advanced lung adenocarcinoma: a retrospective study of patients treated with platinum-based chemotherapy. Eur J Cancer 50, 1819-1828.

26. D'Haene, N., Fontanges, Q., De Neve, N., Blanchard, O., Melendez, B., Delos, M., Dehou, M. F., Maris, C., Nagy, N., Rousseau, E., Vandenhove, J., Gilles, A., De

Prez, C., Verset, L., Van Craynest, M. P., Demetter, P., Van Laethem, J. L., Salmon, I.,Le Mercier, M. (2018): Clinical application of targeted next-generation sequencing for colorectal cancer patients: a multicentric Belgian experience. Oncotarget 9, 20761-20768.

27. D'Haene, N., Le Mercier, M., De Neve, N., Blanchard, O., Delaunoy, M., El Housni, H., Dessars, B., Heimann, P., Remmelink, M., Demetter, P., Tejpar, S.,Salmon, I. (2015): Clinical Validation of Targeted Next Generation Sequencing for Colon and Lung Cancers. PLoS One 10, e0138245.

28. De Roock, W., De Vriendt, V., Normanno, N., Ciardiello, F.,Tejpar, S. (2011): KRAS, BRAF, PIK3CA, and PTEN mutations: implications for targeted therapies in metastatic colorectal cancer. Lancet Oncol 12, 594-603.

29. Donehower, L. A., Creighton, C. J., Schultz, N., Shinbrot, E., Chang, K., Gunaratne, P. H., Muzny, D., Sander, C., Hamilton, S. R., Gibbs, R. A.,Wheeler, D. (2013): MLH1-silenced and non-silenced subgroups of hypermutated colorectal carcinomas have distinct mutational landscapes. J Pathol 229, 99-110.

30. Duman-Scheel, M. (2012): Deleted in Colorectal Cancer (DCC) pathfinding: axon guidance gene finally turned tumor suppressor. Curr Drug Targets 13, 1445-1453.

31. Fearon, E. R.,Vogelstein, B. (1990): A genetic model for colorectal tumorigenesis. Cell 61, 759-767.

32. Ferri, F. F. (2019): Ferri's Clinical Advisor 2020 (1): Elsevier (Elsevier - Health Sciences Division)

33. Fleming, M., Ravula, S., Tatishchev, S. F.,Wang, H. L. (2012): Colorectal carcinoma: Pathologic aspects. J Gastrointest Oncol 3, 153-173.

34. Fleming, N. I., Jorissen, R. N., Mouradov, D., Christie, M., Sakthianandeswaren, A., Palmieri, M., Day, F., Li, S., Tsui, C., Lipton, L., Desai, J., Jones, I. T., McLaughlin, S., Ward, R. L., Hawkins, N. J., Ruszkiewicz, A. R., Moore, J., Zhu, H. J., Mariadason, J. M., Burgess, A. W., Busam, D., Zhao, Q., Strausberg, R. L., Gibbs, P.,Sieber, O. M. (2013): SMAD2, SMAD3 and SMAD4 mutations in colorectal cancer. Cancer Res 73, 725-735.

35. Francipane, M. G.,Lagasse, E. (2014): mTOR pathway in colorectal cancer: an update. Oncotarget 5, 49-66.

36. Fumagalli, D., Gavin, P. G., Taniyama, Y., Kim, S. I., Choi, H. J., Paik, S.,Pogue-Geile, K. L. (2010): A rapid, sensitive, reproducible and cost-effective method for mutation profiling of colon cancer and metastatic lymph nodes. BMC Cancer 10, 101.

37. Galon, J., Mlecnik, B., Bindea, G., Angell, H. K., Berger, A., Lagorce, C., Lugli, A., Zlobec, I., Hartmann, A., Bifulco, C., Nagtegaal, I. D., Palmqvist, R., Masucci, G. V., Botti, G., Tatangelo, F., Delrio, P., Maio, M., Laghi, L., Grizzi, F., Asslaber, M., D'Arrigo, C., Vidal-Vanaclocha, F., Zavadova, E., Chouchane, L., Ohashi, P. S., Hafezi-Bakhtiari, S., Wouters, B. G., Roehrl, M., Nguyen, L., Kawakami, Y., Hazama, S., Okuno, K., Ogino, S., Gibbs, P., Waring, P., Sato, N., Torigoe, T., Itoh, K., Patel, P. S., Shukla, S. N., Wang, Y., Kopetz, S., Sinicrope, F. A., Scripcariu, V., Ascierto, P. A., Marincola, F. M., Fox, B. A.,Pages, F. (2014): Towards the introduction of the 'Immunoscore' in the classification of malignant tumours. J Pathol 232, 199-209.

38. Garcia-Romero, N., Palacin-Aliana, I., Madurga, R., Carrion-Navarro, J., Esteban-Rubio, S., Jimenez, B., Collazo, A., Perez-Rodriguez, F., Ortiz de Mendivil, A.,

Fernandez-Carballal, C., Garcia-Duque, S., Diamantopoulos-Fernandez, J., Belda-Iniesta, C., Prat-Acin, R., Sanchez-Gomez, P., Calvo, E.,Ayuso-Sacido, A. (2020): Bevacizumab dose adjustment to improve clinical outcomes of glioblastoma. BMC Med 18, 142.

39. Genomes Project, C., Abecasis, G. R., Altshuler, D., Auton, A., Brooks, L. D., Durbin, R. M., Gibbs, R. A., Hurles, M. E.,McVean, G. A. (2010): A map of human genome variation from population-scale sequencing. Nature 467, 1061-1073.

40. Gulilat, M., Lamb, T., Teft, W. A., Wang, J., Dron, J. S., Robinson, J. F., Tirona, R. G., Hegele, R. A., Kim, R. B.,Schwarz, U. I. (2019): Targeted next generation sequencing as a tool for precision medicine. BMC Med Genomics 12, 81.

41. Hamada, T., Nowak, J. A.,Ogino, S. (2017): PIK3CA mutation and colorectal cancer precision medicine. Oncotarget 8, 22305-22306.

42. Hamilton, G.,Rath, B. (2018): Circulating Tumor Cells in the Parallel Invasion Model Supporting Early Metastasis. Oncomedicine 3, 15-27.

43. Hanahan, D.,Weinberg, R. A. (2011): Hallmarks of cancer: the next generation. Cell 144, 646-674.

44. Hasan, M. S., Wu, X.,Zhang, L. (2019): Uncovering missed indels by leveraging unmapped reads. Sci Rep 9, 11093.

45. Heemskerk-Gerritsen, B. A., Rookus, M. A., Aalfs, C. M., Ausems, M. G., Collee, J. M., Jansen, L., Kets, C. M., Keymeulen, K. B., Koppert, L. B., Meijers-Heijboer, H. E., Mooij, T. M., Tollenaar, R. A., Vasen, H. F., Hebon, Hooning, M. J.,Seynaeve, C. (2015): Improved overall survival after contralateral risk-reducing mastectomy in

BRCA1/2 mutation carriers with a history of unilateral breast cancer: a prospective analysis. Int J Cancer 136, 668-677.

46. Herold, G. (2018): Innere Medizin 2019: Herold

47. Hughes, L. A., Melotte, V., de Schrijver, J., de Maat, M., Smit, V. T., Bovee, J. V., French, P. J., van den Brandt, P. A., Schouten, L. J., de Meyer, T., van Criekinge, W., Ahuja, N., Herman, J. G., Weijenberg, M. P.,van Engeland, M. (2013): The CpG island methylator phenotype: what's in a name? Cancer Res 73, 5858-5868.

48. Hurwitz, H., Fehrenbacher, L., Novotny, W., Cartwright, T., Hainsworth, J., Heim, W., Berlin, J., Baron, A., Griffing, S., Holmgren, E., Ferrara, N., Fyfe, G., Rogers, B., Ross, R.,Kabbinavar, F. (2004): Bevacizumab plus irinotecan, fluorouracil, and leucovorin for metastatic colorectal cancer. N Engl J Med 350, 2335-2342.

49. Huyghe, N., Baldin, P.,Van den Eynde, M. (2020): Immunotherapy with immune checkpoint inhibitors in colorectal cancer: what is the future beyond deficient mismatch-repair tumours? Gastroenterol Rep (Oxf) 8, 11-24.

50. International HapMap, C., Altshuler, D. M., Gibbs, R. A., Peltonen, L., Altshuler, D. M., Gibbs, R. A., Peltonen, L., Dermitzakis, E., Schaffner, S. F., Yu, F., Peltonen, L., Dermitzakis, E., Bonnen, P. E., Altshuler, D. M., Gibbs, R. A., de Bakker, P. I., Deloukas, P., Gabriel, S. B., Gwilliam, R., Hunt, S., Inouye, M., Jia, X., Palotie, A., Parkin, M., Whittaker, P., Yu, F., Chang, K., Hawes, A., Lewis, L. R., Ren, Y., Wheeler, D., Gibbs, R. A., Muzny, D. M., Barnes, C., Darvishi, K., Hurles, M., Korn, J. M., Kristiansson, K., Lee, C., McCarrol, S. A., Nemesh, J., Dermitzakis, E., Keinan, A., Montgomery, S. B., Pollack, S., Price, A. L., Soranzo, N., Bonnen, P. E., Gibbs, R. A., Gonzaga-Jauregui, C., Keinan, A., Price, A. L., Yu, F., Anttila, V., Brodeur, W., Daly, M. J., Leslie, S., McVean, G., Moutsianas, L., Nguyen, H., Schaffner, S. F., Zhang, Q., Ghori, M. J., McGinnis, R., McLaren, W., Pollack, S.,

Price, A. L., Schaffner, S. F., Takeuchi, F., Grossman, S. R., Shlyakhter, I., Hostetter, E. B., Sabeti, P. C., Adebamowo, C. A., Foster, M. W., Gordon, D. R., Licinio, J., Manca, M. C., Marshall, P. A., Matsuda, I., Ngare, D., Wang, V. O., Reddy, D., Rotimi, C. N., Royal, C. D., Sharp, R. R., Zeng, C., Brooks, L. D.,McEwen, J. E. (2010): Integrating common and rare genetic variation in diverse human populations. Nature 467, 52-58.

51. Irahara, N., Baba, Y., Nosho, K., Shima, K., Yan, L., Dias-Santagata, D., Iafrate, A. J., Fuchs, C. S., Haigis, K. M.,Ogino, S. (2010): NRAS mutations are rare in colorectal cancer. Diagn Mol Pathol 19, 157-163.

52. Jia, M., Gao, X., Zhang, Y., Hoffmeister, M.,Brenner, H. (2016): Different definitions of CpG island methylator phenotype and outcomes of colorectal cancer: a systematic review. Clin Epigenetics 8, 25.

53. Kim, B. H., Kim, J. M., Kang, G. H., Chang, H. J., Kang, D. W., Kim, J. H., Bae, J. M., Seo, A. N., Park, H. S., Kang, Y. K., Lee, K. H., Cho, M. Y., Do, I. G., Lee, H. S., Chang, H. K., Park, D. Y., Kang, H. J., Sohn, J. H., Chang, M. S., Jung, E. S., Jin, S. Y., Yu, E., Han, H. S.,Kim, Y. W. (2020): Standardized Pathology Report for Colorectal Cancer, 2nd Edition. J Pathol Transl Med 54, 1-19.

54. Kim, R., Schell, M. J., Teer, J. K., Greenawalt, D. M., Yang, M.,Yeatman, T. J. (2015): Co-evolution of somatic variation in primary and metastatic colorectal cancer may expand biopsy indications in the molecular era. PLoS One 10, e0126670.

55. Kim, Y. S., Jeong, H., Choi, J. W., Oh, H. E.,Lee, J. H. (2017): Unique characteristics of ARID1A mutation and protein level in gastric and colorectal cancer: A meta-analysis. Saudi J Gastroenterol 23, 268-274.

56. Klein, C. A. (2009): Parallel progression of primary tumours and metastases. Nat Rev Cancer 9, 302-312.

57. Koscielny, S.,Tubiana, M. (2010): Parallel progression of tumour and metastases. Nat Rev Cancer 10, 156.

58. Kuipers, E. J., Grady, W. M., Lieberman, D., Seufferlein, T., Sung, J. J., Boelens, P. G., van de Velde, C. J.,Watanabe, T. (2015): Colorectal cancer. Nat Rev Dis Primers 1, 15065.

59. Landrum, M. J., Lee, J. M., Benson, M., Brown, G. R., Chao, C., Chitipiralla, S., Gu, B., Hart, J., Hoffman, D., Jang, W., Karapetyan, K., Katz, K., Liu, C., Maddipatla, Z., Malheiro, A., McDaniel, K., Ovetsky, M., Riley, G., Zhou, G., Holmes, J. B., Kattman, B. L.,Maglott, D. R. (2018): ClinVar: improving access to variant interpretations and supporting evidence. Nucleic Acids Res 46, D1062-D1067.

60. Lin, E. I., Tseng, L. H., Gocke, C. D., Reil, S., Le, D. T., Azad, N. S.,Eshleman, J. R. (2015): Mutational profiling of colorectal cancers with microsatellite instability. Oncotarget 6, 42334-42344.

61. Lipsyc, M.,Yaeger, R. (2015): Impact of somatic mutations on patterns of metastasis in colorectal cancer. J Gastrointest Oncol 6, 645-649.

62. Madsen, R. R., Knox, R. G., Pearce, W., Lopez, S., Mahler-Araujo, B., McGranahan, N., Vanhaesebroeck, B.,Semple, R. K. (2019): Oncogenic PIK3CA promotes cellular stemness in an allele dose-dependent manner. Proc Natl Acad Sci U S A 116, 8380-8389.

63. Mairinger, F. D., Werner, R., Flom, E., Schmeller, J., Borchert, S., Wessolly, M., Wohlschlaeger, J., Hager, T., Mairinger, T., Kollmeier, J., Christoph, D. C., Schmid, K. W.,Walter, R. F. H. (2017): miRNA regulation is important for DNA damage repair and recognition in malignant pleural mesothelioma. Virchows Arch 470, 627-637.

64. Malapelle, U., Pisapia, P., Sgariglia, R., Vigliar, E., Biglietto, M., Carlomagno, C., Giuffre, G., Bellevicine, C.,Troncone, G. (2016): Less frequently mutated genes in colorectal cancer: evidences from next-generation sequencing of 653 routine cases. J Clin Pathol 69, 767-771.

65. Marley, A. R.,Nan, H. (2016): Epidemiology of colorectal cancer. Int J Mol Epidemiol Genet 7, 105-114.

66. Moiseyenko, V. M., Moiseyenko, F. V., Yanus, G. A., Kuligina, E. S., Sokolenko, A. P., Bizin, I. V., Kudriavtsev, A. A., Aleksakhina, S. N., Volkov, N. M., Chubenko, V. A., Kozyreva, K. S., Kramchaninov, M. M., Zhuravlev, A. S., Shelekhova, K. V., Pashkov, D. V., Ivantsov, A. O., Venina, A. R., Sokolova, T. N., Preobrazhenskaya, E. V., Mitiushkina, N. V., Togo, A. V., Iyevleva, A. G.,Imyanitov, E. N. (2018): First-Line Cetuximab Monotherapy in KRAS/NRAS/BRAF Mutation-Negative Colorectal Cancer Patients. Clin Drug Investig 38, 553-562.

67. Montel, V., Mose, E. S.,Tarin, D. (2006): Tumor-stromal interactions reciprocally modulate gene expression patterns during carcinogenesis and metastasis. Int J Cancer 119, 251-263.

68. Neophytou, C. M., Kyriakou, T. C.,Papageorgis, P. (2019): Mechanisms of Metastatic Tumor Dormancy and Implications for Cancer Therapy. Int J Mol Sci 20

69. Niedermaier, B., Sak, A., Zernickel, E., Xu, S., Groneberg, M.,Stuschke, M. (2019): Targeting ARID1A-mutant colorectal cancer: depletion of ARID1B increases radiosensitivity and modulates DNA damage response. Sci Rep 9, 18207.

70. Niibe, Y.,Hayakawa, K. (2010): Oligometastases and oligo-recurrence: the new era of cancer therapy. Jpn J Clin Oncol 40, 107-111.

71. Noorbakhsh, J., Kim, H., Namburi, S.,Chuang, J. H. (2018): Distribution-based measures of tumor heterogeneity are sensitive to mutation calling and lack strong clinical predictive power. Sci Rep 8, 11445.

72. Oh, M., McBride, A., Yun, S., Bhattacharjee, S., Slack, M., Martin, J. R., Jeter, J.,Abraham, I. (2018): BRCA1 and BRCA2 Gene Mutations and Colorectal Cancer Risk: Systematic Review and Meta-analysis. J Natl Cancer Inst 110, 1178-1189.

73. Papamichael, D., Audisio, R. A., Glimelius, B., de Gramont, A., Glynne-Jones, R., Haller, D., Kohne, C. H., Rostoft, S., Lemmens, V., Mitry, E., Rutten, H., Sargent, D., Sastre, J., Seymour, M., Starling, N., Van Cutsem, E.,Aapro, M. (2015): Treatment of colorectal cancer in older patients: International Society of Geriatric Oncology (SIOG) consensus recommendations 2013. Ann Oncol 26, 463-476.

74. Porru, M., Pompili, L., Caruso, C., Biroccio, A.,Leonetti, C. (2018): Targeting KRAS in metastatic colorectal cancer: current strategies and emerging opportunities. J Exp Clin Cancer Res 37, 57.

75. Rawla, P., Sunkara, T.,Barsouk, A. (2019): Epidemiology of colorectal cancer: incidence, mortality, survival, and risk factors. Prz Gastroenterol 14, 89-103.

76. Reynolds, I. S., O'Connell, E., Fichtner, M., McNamara, D. A., Kay, E. W., Prehn, J. H. M., Furney, S. J.,Burke, J. P. (2019): Mucinous adenocarcinoma of the colon and rectum: A genomic analysis. J Surg Oncol 120, 1427-1435.

77. Riihimaki, M., Hemminki, A., Sundquist, J.,Hemminki, K. (2016): Patterns of metastasis in colon and rectal cancer. Sci Rep 6, 29765.

78. Rivlin, N., Brosh, R., Oren, M.,Rotter, V. (2011): Mutations in the p53 Tumor Suppressor Gene: Important Milestones at the Various Steps of Tumorigenesis. Genes Cancer 2, 466-474.

79. Rychahou, P. G., Kang, J., Gulhati, P., Doan, H. Q., Chen, L. A., Xiao, S. Y., Chung, D. H.,Evers, B. M. (2008): Akt2 overexpression plays a critical role in the establishment of colorectal cancer metastasis. Proc Natl Acad Sci U S A 105, 20315-20320.

80. Samuels, Y., Wang, Z., Bardelli, A., Silliman, N., Ptak, J., Szabo, S., Yan, H., Gazdar, A., Powell, S. M., Riggins, G. J., Willson, J. K., Markowitz, S., Kinzler, K. W., Vogelstein, B.,Velculescu, V. E. (2004): High frequency of mutations of the PIK3CA gene in human cancers. Science 304, 554.

81. Savas, S.,Skardasi, G. (2018): The SWI/SNF complex subunit genes: Their functions, variations, and links to risk and survival outcomes in human cancers. Crit Rev Oncol Hematol 123, 114-131.

82. Sherry, S. T., Ward, M. H., Kholodov, M., Baker, J., Phan, L., Smigielski, E. M.,Sirotkin, K. (2001): dbSNP: the NCBI database of genetic variation. Nucleic Acids Res 29, 308-311.

83. Sinicrope, F. A.,Sargent, D. J. (2012): Molecular pathways: microsatellite instability in colorectal cancer: prognostic, predictive, and therapeutic implications. Clin Cancer Res 18, 1506-1512.

84. Suzuki, M.,Tarin, D. (2007): Gene expression profiling of human lymph node metastases and matched primary breast carcinomas: clinical implications. Mol Oncol 1, 172-180.

85. Taliano, R. J., LeGolvan, M.,Resnick, M. B. (2013): Immunohistochemistry of colorectal carcinoma: current practice and evolving applications. Hum Pathol 44, 151-163.

86. Tan, I. B., Malik, S., Ramnarayanan, K., McPherson, J. R., Ho, D. L., Suzuki, Y., Ng, S. B., Yan, S., Lim, K. H., Koh, D., Hoe, C. M., Chan, C. Y., Ten, R., Goh, B. K., Chung, A. Y., Tan, J., Chan, C. X., Tay, S. T., Alexander, L., Nagarajan, N., Hillmer, A. M., Tang, C. L., Chua, C., Teh, B. T., Rozen, S.,Tan, P. (2015): High-depth sequencing of over 750 genes supports linear progression of primary tumors and metastases in most patients with liver-limited metastatic colorectal cancer. Genome Biol 16, 32.

87. Tate, J. G., Bamford, S., Jubb, H. C., Sondka, Z., Beare, D. M., Bindal, N., Boutselakis, H., Cole, C. G., Creatore, C., Dawson, E., Fish, P., Harsha, B., Hathaway, C., Jupe, S. C., Kok, C. Y., Noble, K., Ponting, L., Ramshaw, C. C., Rye, C. E., Speedy, H. E., Stefancsik, R., Thompson, S. L., Wang, S., Ward, S., Campbell, P. J.,Forbes, S. A. (2018): COSMIC: the Catalogue Of Somatic Mutations In Cancer. Nucleic Acids Res

88. Testa, U., Pelosi, E.,Castelli, G. (2018): Colorectal cancer: genetic abnormalities, tumor progression, tumor heterogeneity, clonal evolution and tumor-initiating cells. Med Sci (Basel) 6

89. Tuppurainen, K., Makinen, J. M., Junttila, O., Liakka, A., Kyllonen, A. P., Tuominen, H., Karttunen, T. J.,Makinen, M. J. (2005): Morphology and microsatellite instability in sporadic serrated and non-serrated colorectal cancer. J Pathol 207, 285-294.

90. Vatandoust, S., Price, T. J.,Karapetis, C. S. (2015): Colorectal cancer: Metastases to a single organ. World J Gastroenterol 21, 11767-11776.

91. Vogelstein, B.,Kinzler, K. W. (2015): The Path to Cancer --Three Strikes and You're Out. N Engl J Med 373, 1895-1898.

92. Vogelstein, B., Papadopoulos, N., Velculescu, V. E., Zhou, S., Diaz, L. A., Jr.,Kinzler, K. W. (2013): Cancer genome landscapes. Science 339, 1546-1558.

93. Wilson, C. H., McIntyre, R. E., Arends, M. J.,Adams, D. J. (2010): The activating mutation R201C in GNAS promotes intestinal tumourigenesis in Apc(Min/+) mice through activation of Wnt and ERK1/2 MAPK pathways. Oncogene 29, 4567-4575.

94. WM, G. (2005): Epigenetic events in the colorectum and in colon cancer. Biochem Soc Trans. 33(Pt 4), 684-688.

95. Wu, Z., Liu, Z., Ge, W., Shou, J., You, L., Pan, H.,Han, W. (2018): Analysis of potential genes and pathways associated with the colorectal normal mucosa-adenoma-carcinoma sequence. Cancer Med 7, 2555-2566.

96. Yachida, S., Jones, S., Bozic, I., Antal, T., Leary, R., Fu, B., Kamiyama, M., Hruban, R. H., Eshleman, J. R., Nowak, M. A., Velculescu, V. E., Kinzler, K. W., Vogelstein, B.,Iacobuzio-Donahue, C. A. (2010): Distant metastasis occurs late during the genetic evolution of pancreatic cancer. Nature 467, 1114-1117.

97. Yaeger, R., Chatila, W. K., Lipsyc, M. D., Hechtman, J. F., Cercek, A., Sanchez-Vega, F., Jayakumaran, G., Middha, S., Zehir, A., Donoghue, M. T. A., You, D., Viale, A., Kemeny, N., Segal, N. H., Stadler, Z. K., Varghese, A. M., Kundra, R., Gao, J., Syed, A., Hyman, D. M., Vakiani, E., Rosen, N., Taylor, B. S., Ladanyi, M., Berger, M. F., Solit, D. B., Shia, J., Saltz, L.,Schultz, N. (2018): Clinical Sequencing Defines the Genomic Landscape of Metastatic Colorectal Cancer. Cancer Cell 33, 125-136 e123.

98. Yashiro, M. (2014): Ulcerative colitis-associated colorectal cancer. World J Gastroenterol 20, 16389-16397.

99. Zhao, J., Han, Y., Li, J., Chai, R.,Bai, C. (2019): Prognostic value of KRAS/TP53/PIK3CA in non-small cell lung cancer. Oncol Lett 17, 3233-3240.

100. Zhunussova, G., Afonin, G., Abdikerim, S., Jumanov, A., Perfilyeva, A., Kaidarova, D.,Djansugurova, L. (2019): Mutation Spectrum of Cancer-Associated Genes in Patients With Early Onset of Colorectal Cancer. Front Oncol 9, 673.

7. List of Figures

8. List of Tables

9. Abbreviations

ACVR2A Activin A Receptor Type 2A
AKT1 AKT Serine/Threonine Kinase 1
AKT2 AKT Serine/Threonine Kinase 2
ANOVA Analysis of variance
APC Adenomatous polyposis coli
APK3 Adenylyl-sulfate kinase 3
ARID1B AT-Rich Interaction Domain 1B
ARIDIA AT-Rich Interaction Domain 1A
ATM ATM Serine/Threonine Kinase
BAP1 BRCA1 Associated Protein 1
BCLAF1 BCL2 Associated Transcription Factor 1
BRAF B-Raf Proto-Oncogene, Serine/Threonine Kinase
BRCA1 BRCA1 DNA Repair Associated
BRCA2 BRCA2 DNA Repair Associated
CA 19-9 Carbohydrate-Antigen 19-9
CDKN2A Cyclin Dependent Kinase Inhibitor 2A
CDX2 *Homeobox protein CDX-2*
CEA Carcinoembryonic antigen
CIMP CpG island methylator phenotype
CIN chromosome instable, chromosome instable
CK20 *Keratin 20*
CK7 Keratin 7
COSMIC *Catalogue Of Somatic Mutations In Cancer*
CRAF Raf-1 Proto-Oncogene, Serine/Threonine Kinase
CRC Colorectal cancer
CTNNB1 Catenin Beta 1
dbSNP Single Nucleotide Polymorphism Database
DCC DCC Netrin 1 Receptor
dMMR mismatch repair deficient
DNA Deoxyribonucleic acid
EGFR Epidermal Growth Factor Receptor, epidermal growth factor receptor
ERBB2 Erb-B2 Receptor Tyrosine Kinase 2
FBXW7 F-Box And WD Repeat Domain Containing 7, E3 Ubiquitin
FFPE formalin-fixed, paraffin-embedded tissue
FGFR1 Fibroblast Growth Factor Receptor 1
FGFR2 Fibroblast Growth Factor Receptor 2
FOLFIRI 5-fluorouracil with leucovorin in combination with irinotecan
FOLFOX oxaliplatin
FOLFOXIRI 5-fluorouracil, leucovorin, irinotecan and oxaliplatin
GNA11 G Protein Subunit Alpha 11
GNAQ G Protein Subunit Alpha Q
GNAS GNAS Complex Locus
GPA33 Cell surface A33 antigen
HNPCC hereditary nonpolyposis colorectal cancer
IDH1 Isocitrate Dehydrogenase (NADP(+)) 1
IDH2 Isocitrate Dehydrogenase (NADP(+)) 2
KDM6A Lysine Demethylase 6A
KEGG Kyoto Encyclopedia of Genes and Genomes
KIT KIT Proto-Oncogene, Receptor Tyrosine Kinase

KRAS KRAS proto-oncogene, GTPase

MAP2K1 Mitogen-Activated Protein Kinase Kinase 1

MAP2K2 Mitogen-Activated Protein Kinase Kinase 2

MAPK1 Mitogen-Activated Protein Kinase 1

MDM2 MDM2 Proto-Oncogene, E3 Ubiquitin Protein Ligase

MET MET Proto-Oncogene, Receptor Tyrosine Kinase

MLH1 MutL Homolog 1, Colon Cancer, Nonpolyposis Type 2

MMR DNA mismatch repair

MSH2 DNA Mismatch Repair Protein Msh2

MSH3 MutS Homolog 3

MSH6 MutS Homolog 6

MSI Microsatellite instability

MSI-H microsatellite instability high

MSKCC Memorial Sloan Kettering Cancer Center

MUC1 Mucin 1

MUC2 Mucin 2

MYC MYC Proto-Oncogene, BHLH Transcription Factor

NF1 Neurofibromin 1

NGS Next generation sequencing

NOS not otherwise specified

NRAS Neuroblastoma RAS Viral (V-Ras) Oncogene Homolog

PALB2 Partner And Localizer Of BRCA2

PBRM1 Polybromo 1

PCR Polymerase chain reaction

PD-1 programmed cell death protein 1

PDGFRa Platelet Derived Growth Factor Receptor Alpha

PDGFRB Platelet Derived Growth Factor Receptor Beta

PD-L1 target programmed death-ligand 1

PIK3CA phosphatidylinositol-4,5-bisphosphate 3-kinase, catalytic subunit alpha

Plks Polo-like kinases

PMS2 PMS1 homolog 2, mismatch repair system component

POLO Polo Like Kinase 1

PRKAR1A rotein Kinase CAMP-Dependent Type I Regulatory Subunit Alpha

PTEN Phosphatase And Tensin Homolog

RCRC right sided colorectal cancer, right sided colorectal cancer

RET Proto-Oncogene Tyrosine-Protein Kinase Receptor Ret

RNF43 Ring Finger Protein 43

RPA1 Replication Protein A1

SF3B1 Splicing Factor 3b Subunit 1

SMAD2 SMAD Family Member 2

SMAD4 SMAD Family Member 4

SMARCA2 SWI/SNF Related, Matrix Associated, Actin Dependent Regulator Of Chromatin, Subfamily A, Member 2

SMARCA4 SWI/SNF Related, Matrix Associated, Actin Dependent Regulator Of Chromatin, Subfamily A, Member 4

SMARCB1 SWI/SNF Related, Matrix Associated, Actin Dependent Regulator Of Chromatin, Subfamily B, Member 1

SOX9 SRY-Box Transcription Factor 9

STK11 Serine/Threonine Kinase 11

TAG-72 Tumor-associated glycoprotein 72

TCF7L2 Transcription Factor 7 Like 2

9. Abbreviations

TGF-β Transforming Growth Factor Beta

TGFβR2 Transforming growth factor, beta receptor

TNM tumor node metastasis

TP53 Tumor Protein P53

TSC1 TSC Complex Subunit 1

TSC2 TSC Complex Subunit 2

UICC classification of the Union internationale contre le cancer

VEGFA vascular endothelial growth factor A

WHO World Health Organization

WT wild-type

XELOX capecitabine and oxaliplatin

10. Acknowledgements

During this dissertation I have received a lot of support and assistance.

First of all, I would like to express my gratitude to my supervisor Prof. Dr. Kurt Werner Schmid for the opportunity I was given to conduct my research at the department of pathology of the Essen University Hospital.

I gratefully acknowledge PD Dr. Fabian Mairinger and Dr. Balazs Hegedüs whose expertise and guidance were invaluable throughout my time as a PhD student.

Furthermore, I would like to thank Dr. Thomas Herold for providing essential data, interesting feedback and insightful comments during our cooperation.

My sincere thanks also go to the experienced pathologists Thomas Hager, MD and Agnes Bilecz, MD for their scientific advice and knowledge in this research field.

I would also like to acknowledge my colleagues Michael Wessolly, Sabrina Borchert, Dr. Robert Walter and Elena Mairinger for their wonderful collaboration.

Finally, I would like to thank and express my gratitude to my wife Dr. Theopisti Maimari for her wise counsel and sympathetic ear, and also for supporting and tolerating me during my thesis.

11. Curriculum Vitae

Der Lebenslauf ist in der Online-Version aus Gründen des Datenschutzes nicht enthalten.