Charles University in Prague
The First Faculty of Medicine

Study program: Biomedicine
Study field: Molecular and Cell Biology, Genetics and Virology

UNIVERZITA KARLOVA
1. lékařská fakulta

MSc. Kateřina Jirásková

Genetic variability in sporadic colorectal cancer:
Searching for novel risk, prognostic and predictive biomarkers

Genetická variabilita u sporadické formy kolorektálního karcinomu:
Hledání nových diagnostických, prognostických a prediktivních biomarkerů

Doctoral Thesis

Thesis Supervisor: Pavel Vodička, MD, PhD.

Thesis Consultant: Veronika Vymetálková, MSc, PhD.

PRAGUE, 2019

Enclosure:

I hereby declare that I wrote this thesis independently and that I have properly cited and quoted all sources and literature used. I further declare that the thesis has not been used to obtain any other academic degree.

I agree with the permanent deposition of the electronic version of the thesis in the database Theses.cz for the purpose of systematical control of similarities between theses.

Prohlášení

Prohlašuji, že jsem závěrečnou práci zpracovala samostatně a že jsem řádně uvedla a citovala všechny použité prameny a literaturu. Současně prohlašuji, že práce nebyla využita k získání jiného nebo stejného titulu.

Souhlasím s trvalým uložením elektronické verze mé práce v databázi systému meziuniverzitního projektu Theses.cz za účelem soustavné kontroly podobnosti kvalifikačních prací.

In Prague/V Praze

Kateřina Jirásková

Identification record:

JIRÁSKOVÁ, Kateřina. *Genetic variability in sporadic colorectal cancer: Searching for novel risk, prognostic and predictive biomarkers [Genetická variabilita u sporadické formy kolorektálního karcinomu: hledání nových diagnostických, prognostických a prediktivních markerů]*. Prague, 2019. 75 pages and 7 attached manuscripts *in extenso*. Dissertation thesis. Charles University in Prague, The First Faculty of Medicine, Institute of Experimental Medicine CAS. Thesis Supervisor MD. Vodička Pavel, PhD.

Identifikační záznam:

JIRÁSKOVÁ, Kateřina. *Genetická variabilita u sporadické formy kolorektálního karcinomu: hledání nových diagnostických, prognostických a prediktivních markerů [Genetic variability in sporadic colorectal cancer: Searching for novel risk, prognostic and predictive biomarkers]*. Praha, 2019. 75 stran a 7 příloh. Dizertační práce. Univerzita Karlova, 1. lékařská fakulta, Ústav experimentální medicíny AV ČR, v. v. i. Školitel MUDr. Pavel Vodička, CSc.

ACKNOWLEDGEMENT

I would like express my gratitude to my supervisor MD Pavel Vodička PhD and to my consultant MSc. Veronika Vymetálková PhD for their guidance and support during my studies.

I would like to thank to Dr. David Hughes and all other collaborative departments for their excellent scientific competence, useful constructive criticism, practical advices and interest in my work.

I am also thankful to my co-worker and friend Soňa Vodenková and to all colleagues from Department of Molecular Biology of Cancer for creating a friendly and stimulating environment in the lab, for their support and (not only) scientific discussions.

The great thanks belong to my family and friends for their patience, love, encouragements and endless support during the entire time of my doctoral studies.

This study was supported by the Grant Agency of Charles University (GAUK 112515), the Internal Grant Agency of the Ministry of Health of the Czech Republic (IGA NT 13424 and AZV MZ 15–26535A) and the Grant Agency of the Czech Republic (GACR 15–14789S and GACR 17-16857S).

SUMMARY

Colorectal cancer (CRC) is a major public health problem worldwide. Despite improvements in the diagnostic process and advancement in the treatment methods, the prognosis remains poor. To improve survival rates, it is important to identify people with the predisposition for CRC and to detect the potentially curable early stage of the disease. Furthermore, identifying those who would have an adverse clinical outcome associated with a particular chemotherapy would help to avoid redundant chemotherapy burden in patients and contribute to enhanced therapeutic efficacy, while minimizing treatment-related toxicity.

The aim of the Thesis was to search for novel promising diagnostic, prognostic and predictive DNA-based biomarkers of sporadic form of CRC. As each patient is genetically unique, these biomarkers would aid clinicians in better diagnosis and/or in the selection of an optimal type of therapy for an individual CRC patient based on their molecular profile. In order to explore this issue, we investigated several candidate genes in healthy individuals as well as in newly diagnosed cancer patients.

The major outcomes of this PhD study, which were fully reported in seven publications included in the present Thesis, are 1) The observation of several candidate single nucleotide polymorphisms in microRNA target regions (miRSNPs) of double strand break repair genes, genes important for CRC etiology and mucin genes to be related either to CRC risk or to clinical outcome, 2) Evidence that miRSNPs in target genes modulate the efficiency of corresponding protein expression, 3) The revelation of genetic variants in NOD-like receptor (NLR) genes contribution to CRC onset and progression of the disease, 4) The identification of the association of several potential functional genetic variants in DNA repair genes with CRC.

Taken together, these studies suggested several novel potential biomarkers for clinical use. However, further studies in independent populations are needed to confirm their clinical significance and to decipher the biologic mechanisms underlying the associations.

Keywords: Colorectal cancer, biomarker, SNP, miRSNP, prognosis, chemotherapy response, DNA repair, CRC pathogenesis, mucin genes, NLR genes

SOUHRN

Rakovina tlustého střeva a konečníku (kolorektální karcinom, KRK) představuje celosvětově závažný zdravotní problém. I přes pokroky v diagnostice a v léčebných metodách zůstává prognóza onemocnění špatná. Pro zlepšení celkové míry přežívání je důležité umět rozpoznat jedince s vyšším rizikem vzniku KRK a odhalit onemocnění v rané potenciálně léčitelné fázi. Současně identifikace pacientů, kteří budou reagovat negativně na konkrétní léčbu, by přispěla ke snížení nadbytečné chemoterapie a k minimalizaci toxicity související s léčbou.

Cílem této práce bylo hledání nových diagnostických, prognostických a prediktivních DNA-biomarkerů pro sporadickou formu KRK. Každý člověk je geneticky jedinečný a nalezení těchto biomarkerů by lékařům usnadnilo diagnózu a výběr optimální terapie pro každého pacienta s KRK na základě jejich molekulárního profilu. Pro dosažení tohoto cíle jsme zkoumali několik kandidátních genů u zdravých jedinců i u nově diagnostikovaných pacientů se sporadickou formou KRK.

Výsledky této PhD práce byly shrnuty v sedmi impaktovaných publikacích. Hlavními závěry jsou: 1) Genetické varianty v cílových oblastech pro vazbu microRNA (miRSNPs) v genech opravy dvouřetězcových zlomů, genech důležitých pro etiologii KRK a mucinových genech souvisí buď s rizikem KRK nebo s odpovědí na léčbu, 2) miRSNPs v cílových genech ovlivňují účinnost exprese odpovídajícího proteinu, 3) Genetické varianty NOD-like receptorů (NLR) přispívají ke vzniku a progresi onemocnění KRK, 4) Funkční genetické varianty v DNA opravných genech jsou asociovány s KRK.

Závěrem, tato disertační práce navrhuje několik nových biomarkerů pro klinické využití. Pro potvrzení klinického významu těchto biomarkerů jsou však nezbytné další studie na nezávislých populacích a porozumění s nimi spojených biologických mechanizmů.

Klíčová slova: rakovina tlustého střeva a konečníku, biomarker, genetické varianty, miRSNP, prognóza, odpověď na léčbu, DNA reparace, patogeneze KRK, mucinové geny, NLR geny

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

2D	Two-dimensional
3D	Three-dimensional
3′UTR	3′-untranslated region
5-FU	5-fluorouracil
BER	Base excision repair
BMI	Body mass index
CART	Classification and regression tree analysis
cfDNA	Cell-free DNA
CIs	Confidence intervals
CRC	Colorectal cancer
CT	Computed tomography
CTC	Circulating tumor cells
DSB	Double strand break
EFS	Event-free survival
EGFR	Epidermal growth factor receptor
FAP	Familial adenomatous polyposis
FIT	Fecal immunochemical test
FOBT	Fecal occult blood test
GWAS	Genome-wide association study
HNPCC	Hereditary non-polyposis colorectal cancer (Lynch syndrome)
HR	Homologous recombination
HRs	Hazard ratios
IFNγ	Interferon gamma
IgE	Immunoglobulin E
LD	Linkage disequilibrium
MAF	Minor allele frequency
MHC	Major histocompatibility complex
miRNA	microRNA
miRSNP	SNP in microRNA binding site
MMR	Mismatch repair
MRN	MRE11-RAD50-NBS1complex

MSI	Microsatellite instability
MSI-H	MSI-high
MSS	Microsatellite stable
MST	Median survival time
NER	Nucleotide excision repair
NHEJ	Non-homologous end-joining
NLR	(NOD)-like receptor
NOD	Nucleotide-binding and oligomerization domain
ORs	Odds ratios
OS	Overall survival
PD-L1	Programmed death-ligand 1
SNP	Single nucleotide polymorphism
TCGA	The Cancer Genome Atlas database
TNM	Tumor–node–metastasis stage system

LIST OF MANUSCRIPTS

This Thesis consists of an overview of research I have been involved in during the time of my PhD studies and which were published from 2016 to 2018. All manuscripts comprise molecular epidemiological studies on different human populations focused on genetic variants in candidate genes.

Manuscript I	Double-strand break repair and colorectal cancer: gene variants within 3′ UTRs and microRNAs binding as modulators of cancer risk and clinical outcome. Naccarati A *et al.* 2016. *Oncotarget* 7(17):23156-69. **IF 5.008**
Manuscript II	MicroRNA-binding site polymorphisms in genes involved in colorectal cancer etiopathogenesis and their impact on disease prognosis. Schneiderova M *et al.* 2017. *Mutagenesis* 32(5):533-542. **IF 2.507**
Manuscript III	Polymorphisms in microRNA binding sites of mucin genes as predictors of clinical outcome in colorectal cancer patients. Vymetalkova V *et al.* 2017. *Carcinogenesis* 38(1):28-39. **IF 4.874**
Manuscript IV	Coding variants in NOD-like receptors: An association study on risk and survival of colorectal cancer. Huhn S *et al.* 2018. *PlosOne* 13(6). **IF 2.806**
Manuscript V	Investigation of single and synergic effects of NLRC5 and PD-L1 variants on the risk of colorectal cancer. Catalano C *et al.* 2018. *PLoS One* 13(2). **IF 2.806**
Manuscript VI	Influence of regulatory NLRC5 variants on colorectal cancer survival and 5-fluorouracil-based chemotherapy. Catalano C *et al.* 2018. *Eur J Gastroenterol Hepatol*. 30(8):838-842. **IF 2.152**
Manuscript VII	Functional polymorphisms in DNA repair genes are associated with sporadic colorectal cancer susceptibility and clinical outcome. Jiraskova K *et al.* 2018. *Int J Mol Sci* 20(1). **IF 3.687**

1. INTRODUCTION

1.1 Colorectal cancer

1.1.1 Incidence and mortality

Colorectal cancer (CRC) is the third most common malignancy and the second leading cause of cancer-related death worldwide with an estimated 1.8 million new cases diagnosed and approximately 881,000 deaths every year (Figure 1) (Bray et al. 2018b). The average lifetime risk for CRC is in the range of 3–5%, meaning that 1 out of 20 persons will develop the disease during their life (Ferlay et al. 2015). It represents a common cancer in both men (3rd most common cancer) and woman (2nd most common cancer) accounting for approximately 10.9% of all cancers in men and 9.5% in women worldwide (Bray et al. 2018b).

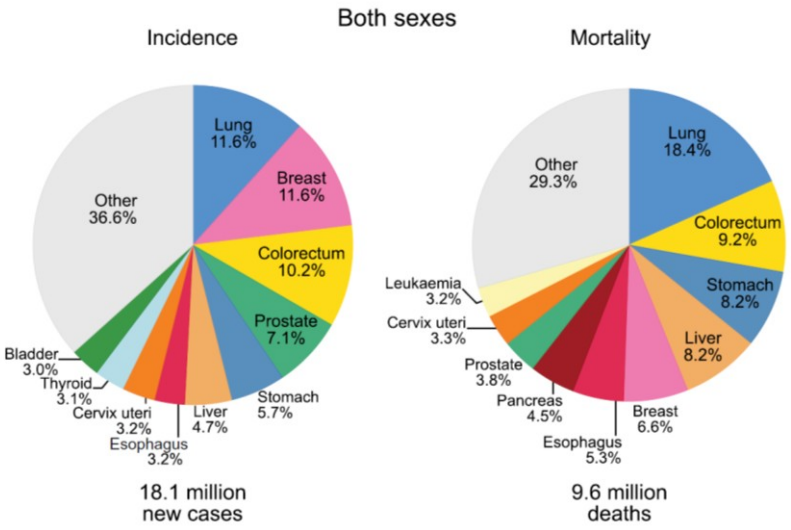


Figure 1: Pie charts represent the incidence and mortality in both sexes of the 10 most common cancers in 2018 (Bray et al. 2018a).

Significant international variations have been observed in CRC distribution with the highest incidence rates in Australia/New Zealand and lowest in Western Africa (Figure 2) (Ferlay et al. 2015; Kuipers et al. 2015). In Europe, estimated CRC incidence rates are highest in Eastern and Central Europe with Czech Republic at the leading ranks (Center et al. 2009; Ferlay et al. 2013).

In the last decades, overall CRC incidence has been stabilizing or declining in western (highly developed) countries: USA, Australia, New Zealand and several European countries (Karim-Kos et al. 2008; Center et al. 2009; Jemal et al. 2010). However, CRC

incidence is still rising in many less developed and economically transitioning countries (low-income and middle-income countries). Increased prevalence may be due to an adoption of a western lifestyle meaning an increased exposure to certain environmental and lifestyle factors such as increased consumption of food rich in sugar and red meat, low physical exercise, obesity and smoking (Center et al. 2009; Arnold et al. 2017; Murphy et al. 2019).

Despite rising incidence in several countries, stabilized or declining trends in CRC mortality have been observed in European countries, North America, and Japan during the last two decades. In Europe, the most favourable trends in mortality were observed in western and northern countries, but there are also declines in mortality in other countries including the Czech Republic (La Vecchia et al. 2010; Bosetti et al. 2011; Hashim et al. 2016; Siegel et al. 2017). Improvement in survival rates are most likely associated with the risk factor reduction (such as smoking, improvements in dietary and lifetime habits), introduction of screening programmes (early diagnosis) and therapeutic improvements (general improvement in surgical techniques for localized tumors as well as new treatment protocols and adjuvant therapies) (Edwards et al. 2010; Fidler et al. 2017).

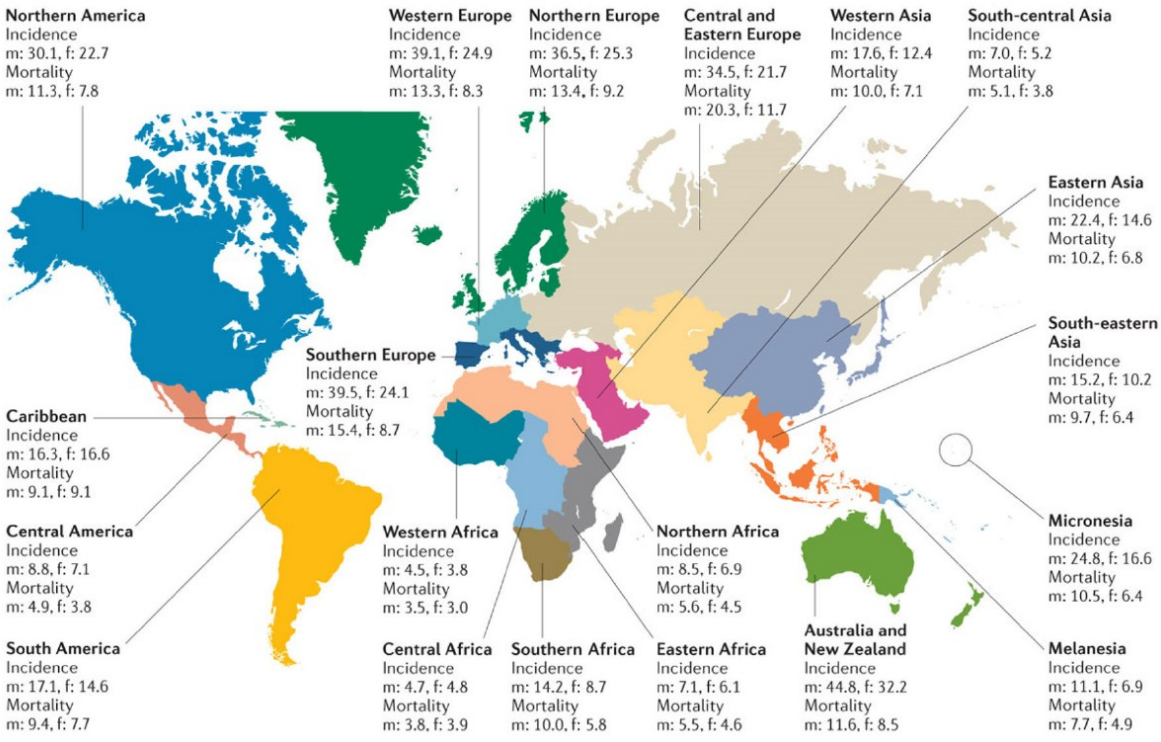


Figure 2: Incidence and mortality rates in males (m) and females (f) (per 100.000 people) across geographic zones (Kuipers et al. 2015).

1.1.2 Risk factors & CRC classification

CRC is a multifactorial disease and the risk of its development could be associated with several factors (lifestyle, socioeconomic, environmental and genetic) (Table 1) (Murphy et al. 2019).

Hereditary conditions account for only a small percentage of all CRCs (about 6% of cases). Highly penetrant germline mutations in known genes, including *APC* in familial adenomatous polyposis (FAP) and DNA mismatch repair genes in hereditary non-polyposis colorectal cancer (HNPCC, Lynch syndrome) are associated with a lifetime risk up to 70-90% (Lynch and de la Chapelle 2003; de la Chapelle 2004).

Familial CRC accounts for up to 20% of cases and comprises patients without an identifiable genetic syndrome but with a family history of CRC (Jasperson et al. 2010; Valle et al. 2019). These cases exhibit common familial risk, likely related to a combination of inherited factors and environment. For individuals with a first degree relative diagnosed at 50–70 years of age the risk of CRC almost doubles. Similarly, if the first degree relative was <50 years of age at diagnosis the risk is three-fold higher than an average risk of the disease. In individuals who have two or more affected family members the risk further increases (Butterworth et al. 2006; Johnson et al. 2013; Kuipers et al. 2015; Samadder et al. 2015).

The majority of CRC cases arise ***sporadically*** (up to 80%) with no specific cause for disease development (Figure 3). However, there are several independent risk factors involved such as age, male sex, diabetes mellitus, previous colonic polyps and intestinal inflammation, dietary and environmental factors (red meat, high-fat diet, obesity, smoking, alcohol consumption, and physical inactivity) and also common low-penetrant genetic variants (Abuli et al. 2010). Independently, these low-penetrant alleles have only a weak effect on the risk of CRC, but in combination they can contribute to a substantial increase in CRC risk, especially when exposed to certain environmental, dietary and lifestyle factors (Goodman et al. 2006; Peters et al. 2015).

Table 1: Summary of the most frequent risk factors associated with a CRC incidence.

Non Modifiable Risk Factors	Modifiable Risk Factors
• Age > 50 • Gender • Inflammatory Bowel Disease (Ulcerative Colitis) • Hereditary Syndromes • Polyps in the colon and/or rectum • Personal/family history of CRC	• Alcohol Consumption • Smoking Tobacco • Diet (rich in red meats, low in fiber) • Obesity • Low Physical Activity

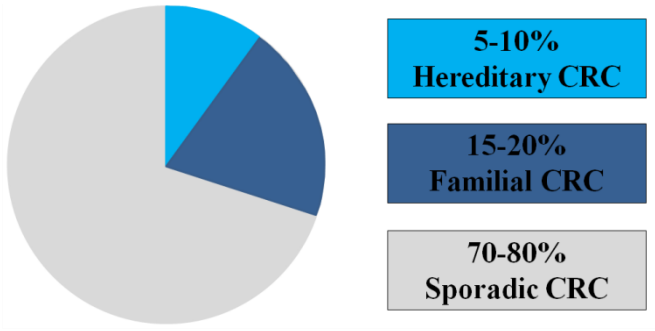


Figure 3: Frequency of sporadic, familial and hereditary forms of colorectal cancer [*modified from Tan 2018*].

1.1.3 Screening

CRC has a long preclinical stage of the disease which offers a large window of opportunity for screening. In individuals with sporadic CRC, the progression from adenomatous polyps to carcinoma takes at least 5–10 years (Fearon 1995; Brenner et al. 2013). Screening can therefore reduce CRC mortality due to the identification of premalignant adenomas (polyps) or detecting potentially curable early stage cancers and thereby preventing the development of the disease by performing an endoscopic removal or surgery (Baxter et al. 2009; Cunningham et al. 2010; Lieberman et al. 2012). Furthermore, the prognosis for patients with CRC is heavily dependent on the stage at diagnosis: 5-year survival is over 90% for patients with early stage cancer, compared with only 5-10% for patients diagnosed with an advanced stage of disease (de la Chapelle 2004; Kuipers et al. 2015).

In most countries, including the Czech Republic, screening is aimed at men and women aged 50–75 years (Suchanek et al. 2018; Schreuders et al. 2015). There are several screening techniques differing in its advantages and limitations (e.g. accuracy, degree of invasiveness, test preparation, required screening interval, and cost) (Table 2) however there is no clear evidence of the superiority of one screening strategy over the others (Stracci et al. 2014; Simon 2016). Patient preference is also an important consideration in the decision-making, however there must be access to follow-up colonoscopy if the clinician recommends it (Wolf et al. 2018).

Table 2: Summary of different screening tests for CRC.

Test	Description	Screening Interval	Advantages	Limitations
Colonoscopy	A procedure allows doctor to look inside the entire colon and rectum with a thin tube with a camera attached to it.	Every 10 years	● High sensitivity (95%) ● Examines entire colon ● Detection of polyps ● Removal of polyps at time of detection	○ Performed at hospital ○ Invasive ○ Cleansing of the colon with laxative ○ Risk of complications
Sigmoidoscopy	Lower part of the colon and rectum are viewed by the doctor with a sigmoidoscope.	Every 5 years combined with FOBT	● High sensitivity (95%) ● Examines entire rectum and 1/2 of the colon ● Removal of polyps at time of detection	○ Requires special facilities ○ Semi-invasive ○ Cleansing of the colon with laxative ○ Screens only distal colon ○ Safety concerns
CT colonography	Uses CT to create 2D and 3D views of the inside of the colon/rectum to detect polyps.	Every 5 years	● High sensitivity (90%) ● Visualization of entire colon ● Detection of polyps	○ Requires special facilities ○ Semi-invasive ○ Cleansing of the colon with laxative ○ Cannot remove lesions at time of detection ○ Radiological safety concerns
FOBT / FIT	Designed to detect occult blood in the stool, which may indicate colon cancer.	Annually	● Done at home ● Noninvasive ● Safe & available	○ Sensitivity FOBT 33%-75% & FIT 60%-80% ○ Poor detection of precancerous lesions ○ Possible false positive test result ○ Cannot remove lesions at time of detection ○ When the test is positive colonoscopy is required

CRC, colorectal cancer; CT, computed tomography; 2D, two-dimensional; 3D, three-dimensional; FOBT, fecal occult blood test; FIT, fecal immunochemical test

1.1.4 Diagnosis & Treatment

CRC diagnosis results either from screening or as a result of an assessment of a patient presenting symptoms such as blood in stool, change in bowel habits, abdominal pain or even a weight loss and fatigue. In symptomatic patients, colonoscopy is the preferred diagnostic method.

Once the disease is diagnosed, current practice to choose and implement the therapy for CRC patients is primarily based on the results of tumor histopathological examination (biopsy) and tumor-node-metastasis (TNM) staging. Treatment regimen substantially differs for colon and rectal cancer patients.

For colon cancer, the cornerstone of treatment is surgical resection and for early stage cancers (I and II), surgery alone may cure the disease. Unfortunately, more than 50% of cases are diagnosed at the higher stage of CRC (III and IV) and the only improvement of the prognosis can be achieved by appropriate 5-fluorouracil (5-FU) based adjuvant therapy (deGramont, XELODA, FOLFOX and FOLFIRI regimes) (Andre et al. 2004;

Twelves et al. 2005; Kuipers et al. 2015). Treatment of rectal cancer is more complex. Patients with stages II and III first undergo radiation therapy (usually simultaneously with 5-FU based chemotherapy) to improve local control of the disease and subsequent surgery is applied (Figure 4) (Hoffe et al. 2010). For stage IV in both, colon and rectum cancer patients, targeted agents are used for treatment alongside 5-FU-based chemotherapy (e.g. cetuximab, panitumumab, bevacizumab, and regorafenib).

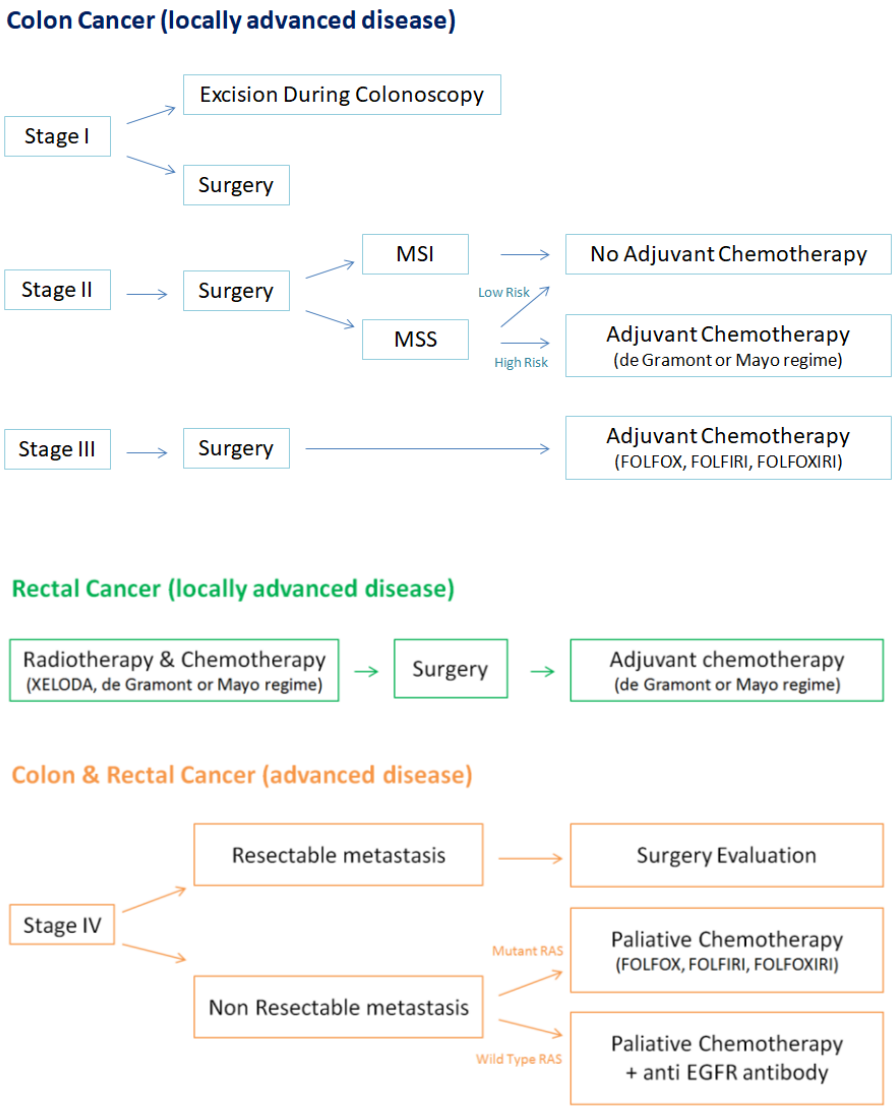


Figure 4: Simplified summary of treatment options for colon and rectal cancer patients [*modified from Aran et al. 2016*].

MSI, microsatellite instable; MSS, microsatellite stable; EGFR, epidermal growth factor receptor

1.2 Biomarkers

A molecular marker (biomarker) is defined as a biological molecule which can be objectively measured and evaluated in blood, and other body fluids, or tissues. Several molecular classes have been studied for their potential use as biomarkers: DNA, cell-free DNA (cfDNA), RNA, microRNA (miRNA), circulating tumor cells (CTC), proteins etc. (Lech et al. 2016; Nikolouzakis et al. 2018). They can be used as an indicator of particular physiological or pathological processes, or pharmacological response to a specified therapeutic intervention (Lee and Chan 2011).

1.2.1 Classification of biomarkers

According to their application, we can categorize biomarkers into three main types as diagnostic, prognostic and predictive (Figure 5) (Gonzalez-Pons and Cruz-Correa 2015; Das et al. 2017; Nikolouzakis et al. 2018).

Diagnostic markers are used to estimate the predisposition for the disease. Therefore according to the risk stratification people at a higher risk of cancer might be recommended for earlier and more intensive screening. Diagnostic markers might be also used for detection of an early stage of the disease. In the case of CRC it means a timely revelation of premalignant polyps.

Prognostic markers give an indication of the clinical outcome (life expectancy) at the time of diagnosis, independent of therapy. They provide information on aggressiveness and the likely progression of the disease including the likelihood of the local recurrence of cancer and/or chance for metastasis.

Predictive markers provide information about the likelihood of the treatment response (benefit of treatment) and outcome parameters such as overall survival (OS) and event-free survival (EFS). As patients exhibit different responses to a certain therapy, markers might help with the choice of the most suitable therapeutic regimen for each patient.

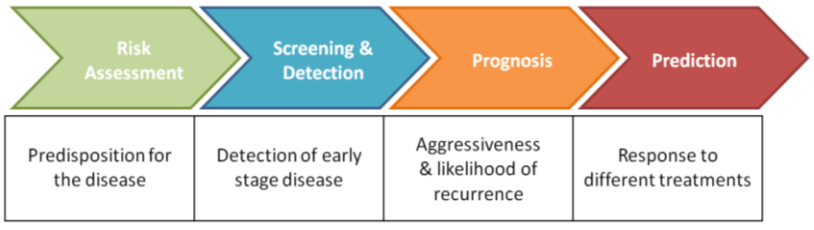


Figure 5: Major types of biomarkers in cancer detection [*modified from https://www.provistadx.com/blog/6-types-of-biomarkers-in-cancer-detection*].

1.2.2 CRC biomarkers in clinical practice

Interestingly, in spite of many published findings on molecular biomarkers in CRC, only a few are nowadays used in daily clinical practice, such as *KRAS*, *BRAF,* and microsatellite instability (MSI).

KRAS mutation status in tumor DNA is analysed in CRC metastatic patients before receiving epidermal growth factor receptor (EGFR) targeted therapy (cetuximab and panitumumab) as only *RAS* wild-type patients benefit from these agents (Kuipers et al. 2015; Vacante et al. 2018). The principle of the therapy is to block the binding of EGF to the EGFR with monoclonal antibodies, which results in blocking the subsequent activation of *RAS* (Figure 6). *RAS* plays a role in a number of intracellular signaling pathways and its dysregulation may ultimately lead to deceased cellular apoptosis, increased cellular proliferation, angiogenesis, and disease metastasis. Mutation in the *RAS* gene causes its constitutively activated GTPase function independent of the binding of EGF to the receptor therefore the targeted therapy does not have any effect. Mutated *RAS* is found in about half of all CRC cases (Coppede et al. 2014; Kocarnik et al. 2015).

BRAF, involved in the same pathway as *KRAS*, is used as a prognostic marker. Mutated *BRAF* is evident in approximately 5-15% of CRC tumors and is associated with a worse survival (Therkildsen et al. 2014; Kocarnik et al. 2015).

MSI status has been shown to be a significant prognostic marker for a better survival and a predictive marker for a worse outcome in terms of response to a standard 5-FU-based chemotherapy, with a trend toward a decreased OS (Bhushan et al. 2009; Coppede et al. 2014; Kocarnik et al. 2015). However, the value of MSI status as a predictive marker for combination chemotherapy regimens (FOLFOX and FOLFIRI) remains uncertain (Kocarnik et al. 2015; Tougeron et al. 2016). MSI is recognized by the presence of increased or decreased number of tandem repeats in microsatellite DNA. High frequency of genetic alterations is caused by mutated or hypermethylated mismatch repair (MMR) genes which results in the inability to correct DNA replication errors. This subsequently leads to a genetic instability and accumulation of DNA errors, both of which may trigger carcinogenesis (Kocarnik et al. 2015). The tumor phenotype associated with this MMR deficiency is reported in approximately 15% of patients with sporadic CRC (Bhushan et al. 2009).

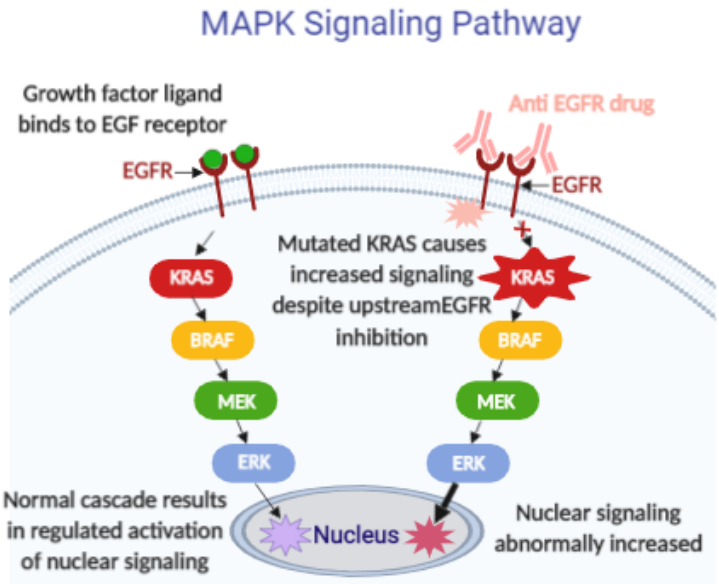


Figure 6: MAPK signaling pathway and anti-EGFR therapy principle [*modified from http://www.apmggroup.net/innovation/molecular_testing/Colon_Pathways/colon.html*].

1.2.3 Novel CRC biomarkers

Despite improvements in our knowledge of the molecular basis of CRC and advancement in the treatment methods, the prognosis remains poor (Figure 7).

CRC is largely asymptomatic until the advanced stage of disease, therefore further progress in diagnostic process is essential to reduce cancer incidence and mortality rates. There are several screening techniques, but they either require a skilled examiner, are invasive to the patient, and involve greater cost (e.g. colonoscopy) or are easy to perform and at reduced cost but less sensitive (fecal occult blood tests) (Schreuders et al. 2015). Since screening is expected to have further impact on CRC management, development of sensitive and specific biomarkers associated with the risk of CRC are being investigated for decades. Little invasive and inexpensive DNA-based tests of blood would be an ideal future possibility.

Furthermore, although advancements in CRC treatment have been made, relapse is still a major factor for the unsatisfactory outcome of the disease. Relapse of CRC after surgical resection with subsequent aduvant chemotherapy, including local recurrence and/or developing metastatic disease, occur in a considerable proportion of these patients (40-50%) within 3 years (Sargent et al. 2007; Gustavsson et al. 2015). Moreover, many CRC patients undergo the systemic chemotherapy without any benefit and even suffer from severe side effects (Longley et al. 2003; Kuipers et al. 2015).

This has provoked a debate over CRC patients at stage II for whom the indication for treatment remains unclear. Chemotherapy based on 5-FU was proved to be beneficial for patients in stage III CRC but the same is not valid for all patients with stage II of the disease (Benson et al. 2004; Quasar Collaborative et al. 2007; Andre et al. 2009; Gangadhar and Schilsky 2010). Chemotherapy can reduce the risk of relapse (20% of patients at stage II CRC will experience recurrence within 5 years) however it can also cause toxicities and impair the quality of patient`s life. Therefore, identification of stage II CRC patients at higher risk of recurrence by molecular markers would help to define those who are likely to benefit from adjuvant therapy and at the same time help to avoid redundant chemotherapy burden in patients at lower risk of recurrence (Lech et al. 2016).

In addition, understanding of different responses to a particular chemotherapeutic agent in patients also remains insufficient. Consequently, as each patient is genetically unique, there is a growing need for novel prognostic and predictive biomarkers. These would aid oncologists in selection of optimal type, combination and dose of drugs for an individual patient to improve the outcome based on their molecular profile. The ultimate goal of a precision medicine approach is to identify patients who would more or less likely benefit from therapy, in other words to contribute to enhanced therapeutic efficacy, while minimizing treatment-related toxicity.

The aim of the current research of biomarkers is to identify and develop highly accurate, non-invasive, rapid and cost-effective biomarkers which could be easily translated into clinical practice. Although prognostic or predictive value of an individual biomarker may be significant, it is likely that a combination of several biomarkers may be utilized into a panel to provide greater information. Many genetic and epigenetic biomarkers have been studied, but still none have been validated for clinical use.

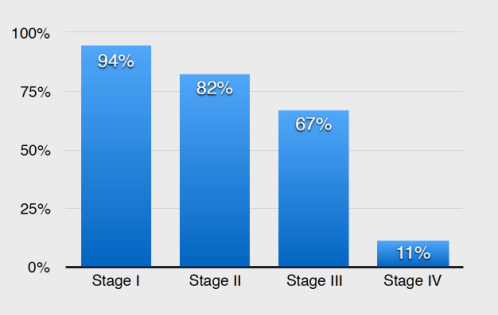


Figure 7: The 5-year survival rates for CRC patients according to a stage of the disease (Lansdorp-Vogelaar et al. 2009).

1.2.3.1 DNA-based biomarkers

DNA, RNA, proteins, metabolites, or processes such as apoptosis, angiogenesis or proliferation have been associated with every type of cancer and therefore might be used as biomarkers. According to the type of molecule used as a biomarker, we can categorize molecular markers in several classes: DNA-based, RNA-based (mRNA, miRNA, piwi-interacting RNA, small interfering RNA, long non-coding RNA etc.), proteins and others.

DNA-based biomarkers include deletions, insertions, loss-of-heterozygosity, MSI, DNA hypermethylation, single nucleotide polymorphisms (SNPs) and other variations on the DNA sequence level (Sidransky 2002).

SNPs are the most frequently studied type of DNA variation. They refer to a substitution of a single nucleotide that occurs at a specific position in the genome, where each variation is presented in > 1% within a population. They are reproducible and can be measured at any point in time (may be used in both prospective and retrospective studies). In most applications SNPs are diallelic, resulting in three possible genotypes (wild type homozygote, heterozygote and variant homozygote).

Genetic variants may be distinguished according to their position in the genome: SNPs may fall within coding sequences of genes, non-coding regions of genes, or regions between genes (intergenic regions) (Figure 8). Polymorphisms within a *coding region* may be further categorized as synonymous and nonsynonymous genetic variants. Synonymous SNPs do not change the amino acid sequence of the final protein due to degeneracy of the genetic code. In the case of nonsynonymous SNPs the amino acid sequence of the protein is changed and the change is classified either as missense variant (single change in the base results in change in amino acid of protein) or nonsense variant (resulting in a premature stop codon).

Non-coding SNPs are located within the gene's regulatory sequences (promoters, enhancers, silencers, and other regulatory regions) and may affect gene splicing, the sequence of non-coding RNA, or timing, location, or level of gene expression. miRNAs binding SNPs also called miRSNPs, located in the 3′-untranslated regions (3′UTR) of genes, are an example of non-coding genetic variants. miRSNPs are able to alter the strength of miRNAs binding to the target mRNA. miRNAs, small non-coding regulatory RNAs, base pair to a complementary motif of the target mRNA. Thus modulating existing binding sites or creating novel binding sites by miRSNPs are

suggested to affect miRNA function and consequently change the expression of target genes (Hammond 2015; Crocco et al. 2016; Slaby 2016; Abba et al. 2017).

Understanding the effects of genetic variants is a difficult process and alteration in the final protein caused by a base change may be indicated as benign, pathogenic, or of unknown significance. Lately, genome-wide association studies (GWAS) have enabled a rapid discovery of SNPs contributing to both disease susceptibility and treatment response by comparing regions of genome between cohorts of patients and healthy controls (Ziegler et al. 2012; Fernandez-Rozadilla et al. 2013a). Regarding CRC, GWAS have been successful in identification of a number of low penetrance SNPs involved in CRC susceptibility however none have still been validated as biomarkers for clinical use (Tomlinson et al. 2011; Dunlop et al. 2012; Fernandez-Rozadilla et al. 2013b; Peters et al. 2013; Huyghe et al. 2019).

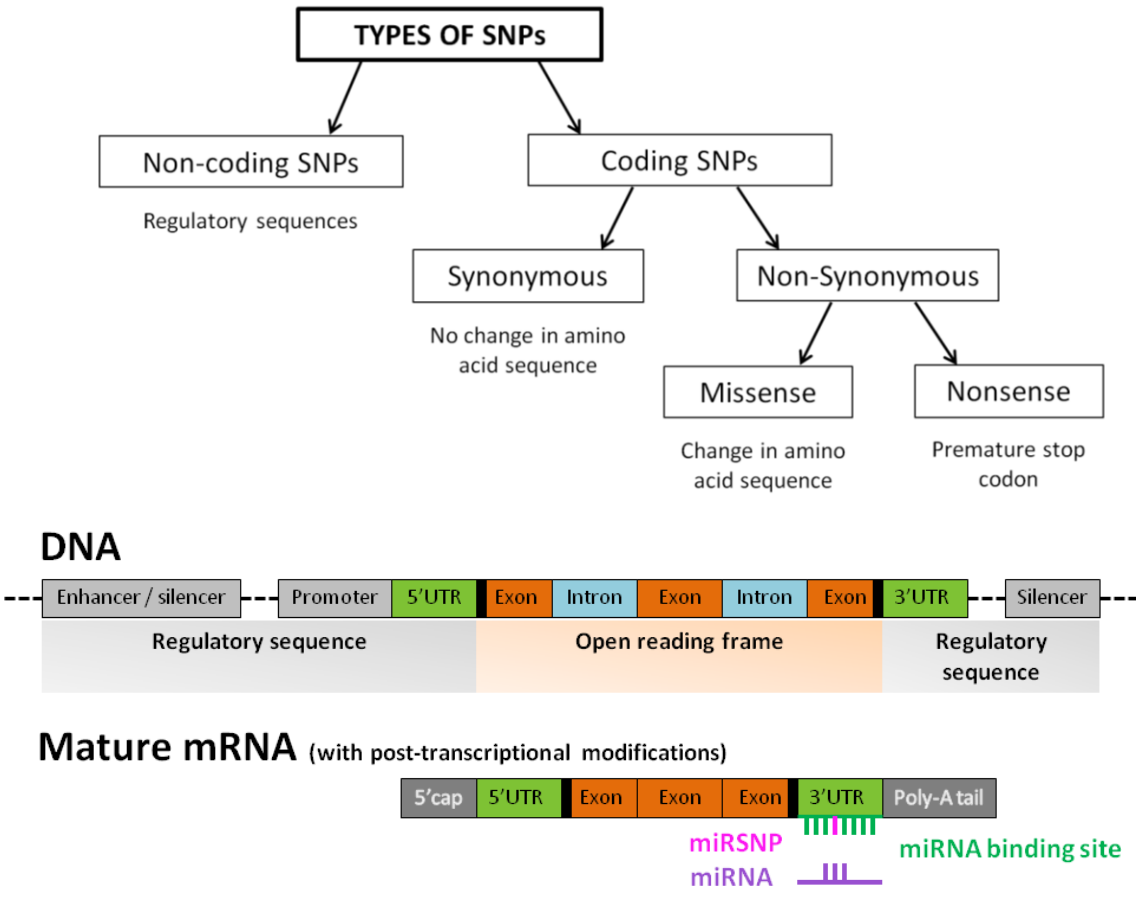


Figure 8: Genetic variants distinguished due to their position in the genome [*modified from Shafee 2017* and *https://en.wikipedia.org/wiki/Single-nucleotide_polymorphism*].

SNP, single nucleotide polymorphism; UTR, untranslated region; miRNA, microRNA; miRSNP, microRNA binding SNP

1.3 Candidate genes for CRC biomarkers

Understanding the genetic basis of the disease has become an important target for research as a better comprehension may lead to an improved prevention or treatment.
There are two main approaches designed to detect associations between genetic factors and the disease course in samples from populations. One is based on a study of candidate genes and the other on testing the entire genome (GWAS) (Amos et al. 2011). Both approaches comprise a combination of benefits and drawbacks.
A candidate gene study is a hypothesis-based approach where the success depends upon the correct choice of genes/pathways to examine, which is exposed to the risk of arbitrariness. However, these studies tend to have higher statistical power than GWA studies that use large numbers of SNPs (Amos et al. 2011).
GWAS is a hypothesis-free approach which can detect genes regardless of whether their function was known before (Cooke et al. 2008). For example, DNA regions important for diabetes or Crohn's disease development have been recognized by GWAS (Sladek et al. 2007; Barrett et al. 2009; Sharp et al. 2015). However, the list of thousands SNPs associated with the susceptibility to complex diseases, identified by GWAS, poses a problem in form of costly validation studies on a large number of individuals.
No conclusion has been reached about which of these two approaches is more effective/convenient. Studies included in this Thesis are of a candidate gene approach and the investigated genes were selected according to published studies on CRC, providing a tremendous amount of information on genes, pathways, and chromosomal regions that appear to be linked to disease. Genes involved in CRC mutagenesis were naturally prime candidates.

1.3.1 DNA repair genes

The human genome is constantly attacked by a plethora of mutagens that impact its stability important for preventing carcinogenesis. A continuous surveillance by DNA repair systems is therefore essential for the maintenance of genome integrity.
Variety of DNA lesions arises from environmental (e.g. ultraviolet component of sunlight, ionizing radiation and numerous genotoxic chemicals including a cigarette smoke) or endogenous (e.g. products of normal cellular metabolism such as reactive oxygen) genotoxic agents. These lesions might interfere with DNA replication, block transcription or in case of double strand breaks (DSBs) are particularly relevant for the

recombination machinery. Unrepaired damages may affect cell metabolism, trigger cell-cycle arrest or contribute to oncogenesis. When damage is too significant, a cell may opt for initiating the apoptosis.

DNA repair mechanisms have evolved in a set of sophisticated, interwoven pathways that recognize and remove different types of DNA lesions in an error-free, or in some cases, error-prone way (Figure 9). The repair machinery operating in mammals has arbitrarily been divided into several pathways including nucleotide excision repair (NER), base excision repair (BER), DSB repair, MMR and direct repair. To allow the cells to repair the DNA damage, different pathways are active throughout different stages of the cell cycle. ***NER*** removes helix-distorting adducts on DNA that interfere with base pairing and generally obstruct transcription and normal replication. Most of these lesions arise from exogenous sources. ***BER*** is responsible for removing small chemical alterations of bases (deamination, oxidation or alkylation) that may impede transcription and replication. BER is mostly concerned with damage arising spontaneously within the cell. Two main pathways are involved in ***DSB*** repair: error-free homologous recombination (HR) dominates in S and G2 phase when the DNA is replicated and error-prone non-homologous end-joining (NHEJ) is most relevant in the G1 phase of the cell cycle. DSBs may arise from ionizing radiation or X-rays, free radicals, chemicals and during replication. As both strands are affected, DSBs are generally considered to be the most deleterious type of DNA damage. ***MMR*** is responsible for removal of mismatched bases erroneously incorporated during replication and deletion/insertion loops within repetitive DNA sequences that have arised from strand slippage during replication or during recombination. ***Direct repair*** is the simplest form of DNA repair with a direct reversal of the lesion as O^6-alkylguanine methyltransferase removes non-native alkyl groups from the guanine residue. These lesions might be induced by dietary nitrosamines or chemotherapy agents. Each of these pathways has been reviewed in depth elsewhere (Hoeijmakers 2001; Christmann et al. 2003; Curtin 2012).

Dysregulation of repair genes is associated with significant health problems and research over the past years show evidence that inherited or acquired deficiencies in DNA repair systems contribute significantly to an increased risk of cancer onset and progression of carcinogenesis, including CRC (Vineis et al. 2009; Curtin 2012).

The importance of proper DNA repair may be also illustrated by the fact that mutations in a number of DNA repair genes are known to be associated with several hereditary

syndromes which are characterised by increased incidence of multiple cancers, immunodeficiency and multiple metabolic alterations (e.g. NER associated Xeroderma pigmentosum and Cockayne's syndrome, DSB associated Bloom syndrome and Fanconi anemia) (Christmann et al. 2003). Concerning the MMR pathway, defects in the genes result in Lynch syndrome, which represents a familial susceptibility to CRC and to a variety of sporadic cancers such as CRC with MSI-H phenotype (Peltomaki 2001; Grady and Markowitz 2015).

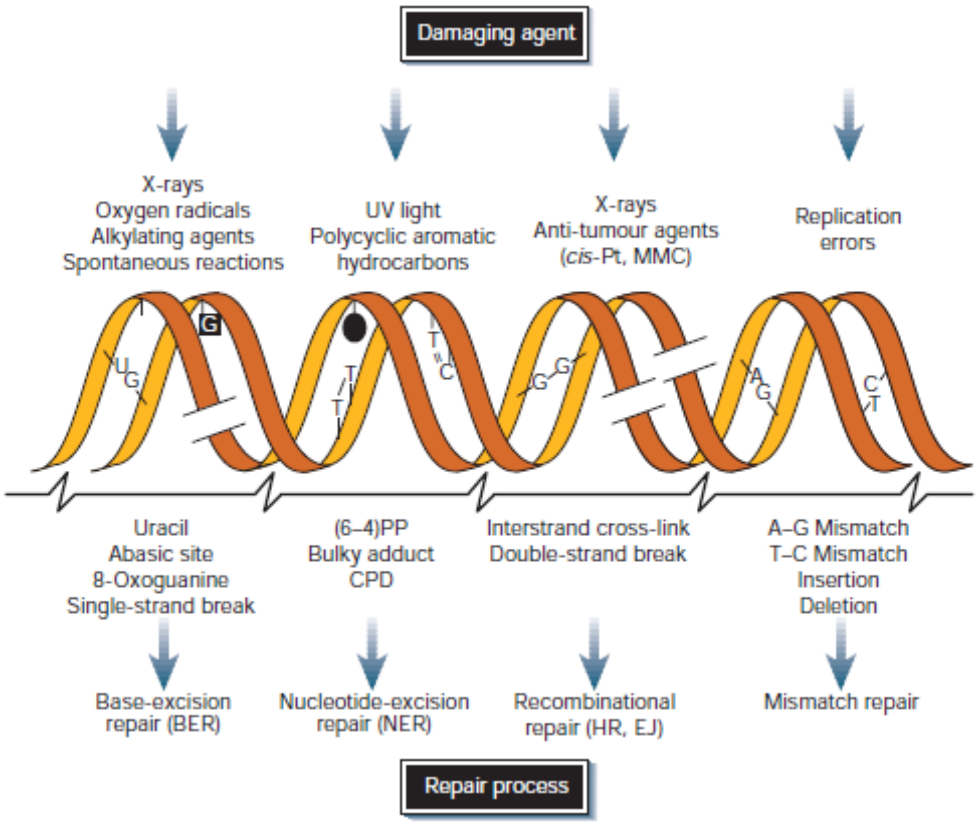


Figure 9: DNA repair pathways (Hoeijmakers 2001).

At least 150 genes have been identified to be associated with DNA repair machinery in humans, many of them being polymorphic in the human population (Friedberg 2003; Roos et al. 2016). As carcinogenesis generally depends on the acquisition of mutations in the cellular DNA, inter-individual differences in DNA repair systems caused by common SNPs in corresponding genes may be expected to play a role in modulating the individual risk of developing cancer or sensitivity of tumor cells to survive DNA damage induced by chemotherapeutic agents. Some SNPs have already been reported as associated with cancer susceptibility in a number of malignancies, including CRC (Naccarati et al. 2007; Pardini et al. 2008; Pardini et al. 2019). However, the

consequences of the majority of SNPs in DNA repair genes have not been fully explored thus far (Xi et al. 2004; Vodicka et al. 2007; Slyskova et al. 2014).

1.3.2 Genes involved in CRC pathogenesis

In 1990, Fearon and Vogelstein described the development of sporadic CRC as a multistep process of accumulation of activating mutations in oncogenes and deactivating mutations in tumor-suppressor genes (Figure 10) (Fearon and Vogelstein 1990; Carethers and Jung 2015). As the intestinal epithelium has a high turnover rate, the constant proliferation in the normal mucosa has to be maintained by the equilibrium between growth promoting oncogenes and growth limiting tumor suppressor genes (Raskov et al. 2014). Thus each of genetic events in these genes confers a selective growth advantage to an affected colon/rectal epithelial cell and may ultimately result in uninhibited cell growth, proliferation, and clonal tumor development (Calvert and Frucht 2002). This adenoma-carcinoma sequence model, which can proceed for more than 10 years, observes a slow development from aberrant crypt proliferation to adenomatous polyps, then to carcinomas *in situ* and finally to malignant tumors (de la Chapelle 2004).

In this model, the inactivation of tumor suppressor gene *APC* by mutations counts as the early step in CRC tumorigenesis and is associated with the initiation of adenoma formation. Larger adenomas and early carcinomas are further promoted in frank carcinoma by acquiring mutations in the *KRAS* oncogene, followed by loss of chromosome 18q with the tumor suppressor gene *SMAD4*, which is downstream of transforming growth factor-β, and mutations in another tumor suppressor gene *TP53* (Brenner et al. 2014). Although cumulative effect of these genetic alterations, rather than their order, determines the biological behavior of the tumor. Nevertheless, *APC* mutations usually occur early in the process and mutations of the *TP53* usually occur late in the process (Calvert and Frucht 2002).

Given that the majority (55–70%) of CRC tumors arises via this pathway, a large number of previous studies focused on the above mentioned genes (Kocarnik et al. 2015). Furthermore, according to this model of CRC carcinogenesis, subtle effects caused by SNPs in genes involved in the CRC pathogenesis may contribute to the disease onset and patient's prognosis.

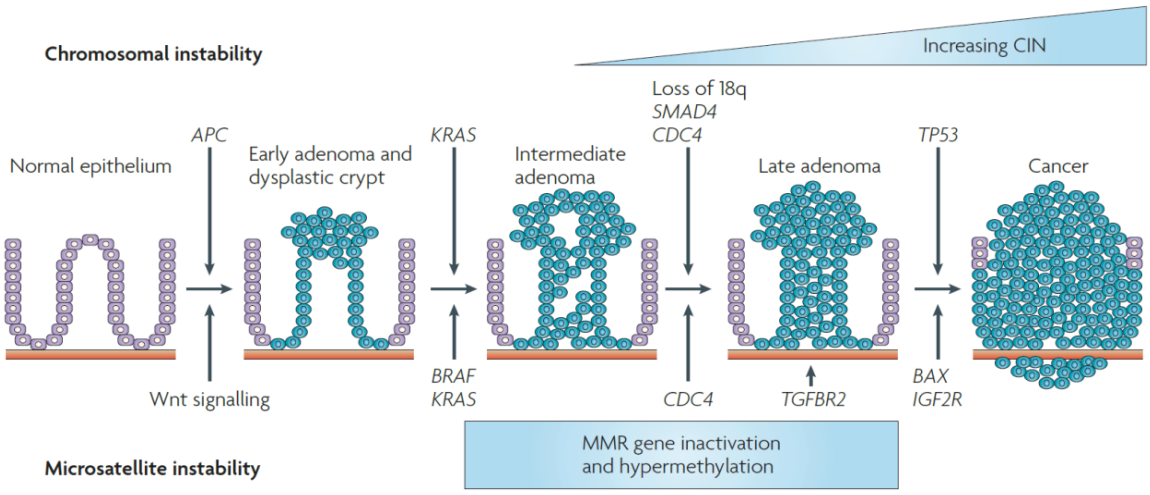


Figure 10: Adenoma-carcinoma sequence model for sporadic CRC (Walther et al. 2009).

1.3.3 Mucin genes

As the cells in contact with the external environment are constantly exposed to ingested toxins, pollutants, luminal contents (that include proteases) and a number of micro-organisms, defense mechanisms, such as the secretion of mucus by mucin genes, have been developed during the evolution.

Mucins are high molecular weight, heavily glycosylated extracellular proteins produced by epithelial cells (Hollingsworth and Swanson 2004; Andrianifahanana et al. 2006). The human mucin family, consisting of members designated *MUC1-MUC21*, includes proteins containing tandem repeat structures with a high proportion of prolines, threonines, and serines. According to the structure and function we may divide mucins in two distinct classes: secreted gel-forming mucins and transmembrane mucins. Together they constitute a mucous barrier (Byrd and Bresalier 2004).

Under physiological conditions, mucins play an important role in the maintenance of homeostasis by covering epithelial surfaces, including human colorectal epithelium, by gel mucus layer (Figure 11) (Gupta et al. 2012). Their function lies in limiting the activation of inflammatory responses at the interface with the environment, therefore deregulation of mucin production is an important link between inflammation and cancer (Kufe 2009).

Aberrant expression of mucins was found in diverse human adenocarcinomas (in gastric, esophageal, breast, endometrial and lung cancer), while increased levels of mucin production have been associated with higher risk of cancer as well as with worse patient prognosis (Yonezawa et al. 2008a; Yonezawa et al. 2008b; Yonezawa et al.

2011; Zheng et al. 2019). Concerning CRC, an overexpression of *MUC1* and *MUC5AC* and a downregulation of *MUC2* have been described to be involved in the development and progression of the disease (Nikolouzakis et al. 2018) and the upregulation of *MUC20* have been observed as a predictor of poor prognosis (Xiao et al. 2013; Lu et al. 2019).

As miRNAs have emerged as important regulators responsible for an altered mucin expression (Macha et al. 2015) miRSNPs located in mucin genes may play a role in cancer susceptibility, efficacy of chemotherapy, and survival.

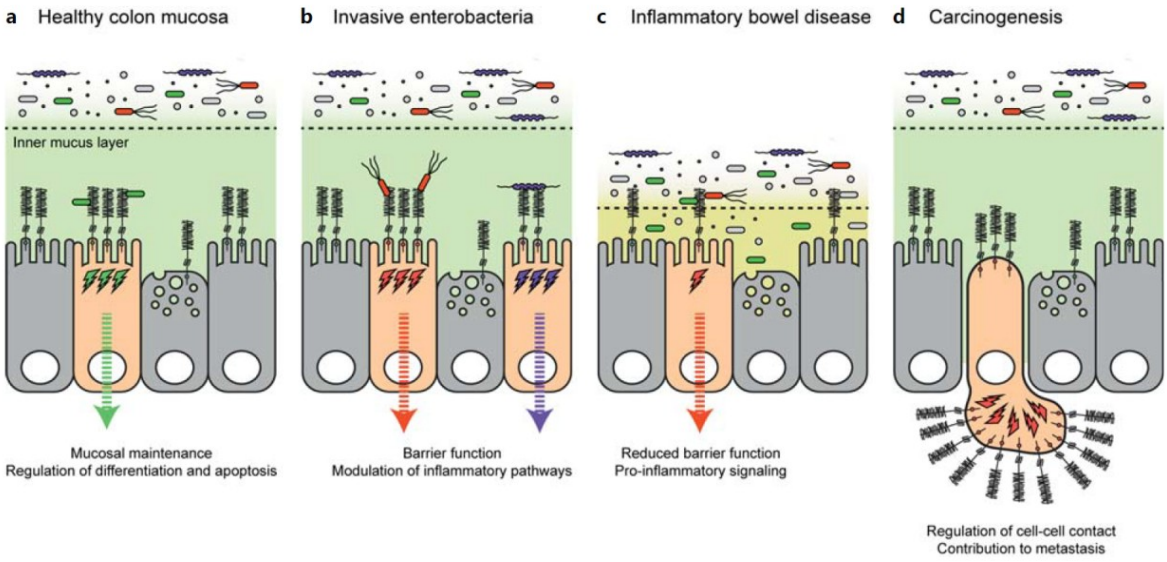


Figure 11: Functions of transmembrane mucins in the human colon (van Putten and Strijbis 2017).

1.3.4 NOD-like receptor genes

Highly conserved nucleotide-binding and oligomerization domain (NOD)-like receptors (NLRs) are cytosolic pattern recognition receptors that play a crucial role in mucosal immune defense (Claes et al. 2015).

As the intestinal tract is continuously interacting with pathogenic or endogenous microorganisms as well as to commensal bacteria, proper immune response and homeostasis between immunity and tolerance has to be strictly controlled. NLRs are activated by recognizing a wide range of pathogens or damage-associated molecular patterns and trigger sequential activation of intracellular signalling pathways that initiate the innate response and the subsequent adaptive immune response (Kim et al. 2016). Therefore a dysregulated expression of NLR genes, due to functional or genetic defects, may lead to an excessive or uncontrolled signalling of underlying regulatory

pathways. Consequently, this may result in the development of local and chronic inflammation, inflammatory bowel disease (such as ulcerative colitis and Crohn's disease) and/or CRC (Belkaid and Hand 2014; Claes et al. 2015).

NLR family comprises of 22 genes in humans and their abnormalities are linked to various diseases (Table 3) (Fukata et al. 2009; Zhong et al. 2013; Kim et al. 2016). A significant association of *Nod2* mutations with a risk of Crohn's disease and SNPs in *Nod1* with the inflammatory bowel disease onset has been already identified by GWAS (Fukata et al. 2009). As the inflammation affects all stages of tumorigenesis, SNPs located in NLRs genes may also play a role in cancer susceptibility.

Table 3: Genetic associations of NLR to inflammatory disorders [*modified from Fukata et al. 2009]*.

NLR	Inflammatory disorder
NOD1	Asthma, atopic eczema, increased serum IgE
NOD2	Asthma, atopic eczema, increased serum IgE, Crohn's disease and Ulcerative colitis

IgE, Immunoglobulin E

2. AIMS

The main goal of the thesis was to identify potential novel biomarkers for sporadic CRC. For this purpose, we studied the association of genetic variants with CRC susceptibility and patient's clinical outcome. We were particularly interested in DNA repair genes, genes involved in CRC pathogenesis (irrespectively of MSI status), mucin genes and cancer-related immunity genes.

We stated the following aims:

- To test whether SNPs in miRNA target regions of selected genes affect cancer susceptibility, survival and response to therapy in CRC patients.

- To examine whether miRSNPs modulate the efficiency of corresponding proteins translation.

- To investigate whether coding SNPs in NLR genes contribute to human CRC development or progression.

- To search whether SNPs causing amino acid substitution in DNA repair genes influence the risk of CRC and modulate the clinical outcome after CRC diagnosis.

3. MATERIAL & METHODS

3.1. Study populations

Manuscript I: Study was conducted on 1126 histologically confirmed CRC patients and 1469 healthy controls from the Czech Republic. All 2595 individuals were interviewed for their lifestyle habits, body mass index (BMI), diabetes, and family/personal history of cancer with a structured questionnaire. Patient's clinical data at the time of diagnosis were collected (location of the tumor, TNM stage and grade) along with information about adjuvant chemotherapy treatment, distant metastasis, relapse and date of death. All subjects were sampled for peripheral blood.

Manuscript II: Study was performed on the same cohort of patients and healthy controls as described in **Manuscript I**. In this case, 1111 CRC patients and 1469 healthy controls, that provided peripheral blood samples, were tested in the study.

Manuscript III: Study was carried out on the same group of patients and healthy controls as in **Manuscript I and II**. Blood samples were collected from 1111 patients and 1469 healthy controls.

Manuscript IV: In this association study, a discovery cohort from the Czech Republic (1237 CRC cases and 787 healthy controls) and replication cohorts from Germany (1798 CRC cases and 1810 healthy controls) and Scotland (2210 CRC cases and 9350 healthy controls) were included. All subjects provided necessary information requested in the questionnaire. Clinical data were collected in collaboration with attending patient's physicians. Blood samples were collected from all study participants.

Manuscript V: In this case-control study, 1424 newly diagnosed patients with sporadic CRC were compared to 1114 age-matched healthy individuals. All study participants were of Czech origin and provided blood samples.

Manuscript VI: A hospital-based study involved 589 incident patients from the Czech Republic diagnosed for sporadic CRC. Clinico-pathological data and information about recurrence, distant metastasis, or date of death were provided. All subjects were sampled for peripheral blood. For expression analyses, patients providing biopsy material and healthy controls providing buffy coats or whole blood were recruited at the University Hospital TuÈbingen.

Manuscript VII: The association study was carried out on a discovery cohort from the Czech Republic (1832 CRC cases and 1172 healthy controls) and replication cohort from Austria (950 CRC cases and 820 healthy controls). Characteristics of both cohort participants were collected. Blood samples were collected from healthy controls. Patients were sampled for nonmalignant colon/rectal tissue or peripheral blood.
All individuals included in the above studies provided informed consent and the particular studies have obtained appropriate approval from Ethic committees.

3.2. Selection of candidate SNPs

To select the candidate SNPs, several *in silico* prediction software programs were implemented (Manuscript I – MicroSNiper, Mirnsnpscore and Polymirt; Manuscript II - MicroSNiPer, miRSNP, Mirnsnpscore, Polymirt, RNAcofold, miRanda, TargetScan, GTEx and SCAN database; Manuscript III - MicroSNiper, Mirnsnpscore and Polymirt; Manuscript IV – SIFT, PolyPhen2, GERP, PhastCons and PhyloP; Manuscript V – Regulome DB, Gtex Portal, MicroSNiper, PERFECTOS-APE and s-TRAP; Manuscript VI - Ensembl, Regulome DB, Gtex Portal, MicroSNiPer and Transcription factor Affinity Prediction; Manuscript VII - F-SNP, GERP, SiPhy, ELASPIC and DUET).
The SNPs were filtered for their minor allele frequency (MAF > 1-10% depending on the Manuscript) in Caucasian populations to reach an appropriate representation of all genotypes in our set of cases and controls (Source: 1000Genomes, dbSNP, HapMap).
SNPs with the required MAF were tested for the possibility to be in linkage disequilibrium (LD) (Source: HaploView and HapMap).
Detailed workflow for the SNPs selection is described in the Manuscripts.

3.3. Genotyping analysis

Genomic DNA was isolated from peripheral blood lymphocytes using standard procedures. When blood was not available, nonmalignant colon/rectal tissue was used to obtain DNA by using the DNeasy Blood and Tissue Kit (Qiagen).
SNPs were determined by TaqMan SNP Genotyping Assays (Thermo Fisher Scientific), KASP Genotyping Assays (LGC genomics), the Illumina HumanCytoSNP or Illumina HumanOmniExpress Platform (Peters et al. 2013), the Infinium Human Exome

BeadChip (Illumina), OmniExpressExome BeadChip (Illumina) and Axiom Genome-Wide CEU 1 Array (Affymetrix). In manuscript IV, imputation was performed for autosomal SNPs to the CEU population in HapMap and in manuscript VII, genotype assignment was performed as described in (Hofer et al. 2017).
The genotypes with unclear results were excluded from the studies.

3.4. Gene expression analysis

In Manuscript I, a Dual-Luciferase reporter assay was used to investigate whether the *MRE11A* rs2155209 alleles were associated with a differential gene expression. The assays were carried out using the Dual-Luciferase Reporter Assay Kit (Promega) and three replicates of all experimental points were performed in each experiment. Luminescence intensity was evaluated by a luminometer (Optima FluoStar) and luciferase activities were averaged from four measurements.
In Manuscipt IV, gene expression analysis was carried out using single-gene TaqMan1Gene Expression Assays (Applied Biosystems). mRNA was isolated from samples of two donors or cell lines (THP-1, HCT116, DLD-1 or CaCo2) by RNeasy Mini Kit (Qiagen). Commercially available RNA samples for human ovary, duodenum, ileum, rectal and colon adenocarcinoma were also used (Agilent). RNA from ileum or colon biopsies was isolated using TRIzol Reagent (Life Technologies) according to standard protocols. The samples were analyzed in triplicate using the 7500fast Real-Time System (Applied Biosystems).

3.5. Statistical analysis

Genotype frequencies in healthy controls were tested for Hardy-Weinberg equilibrium (HWE; Pearson's goodness-of-fit $\chi 2$ test). Odds ratios (ORs) and 95% confidence intervals (CIs) for associations between genotypes and risk of CRC were estimated by logistic regression.
The outcome variables measured were OS (time from diagnosis until death or censorship), and EFS (time from surgery or end of chemotherapy until date of relapse, death or censorship whichever came first). The relative risk of death was estimated as

hazard ratio (HR) using Cox regression. The survival curves for OS and EFS were derived by the Kaplan–Meier method.
In Manuscript I, the multifactor analysis of variance with interactions (MANOVA) was performed in the *in vitro* assays to compare the ratios of the measurements of luminescence between genotypes. In Manuscript V, additive influence of the risk alleles on CRC risk and patient's survival was estimated. In Manuscript VII, a multivariate analysis, referred to as a classification and regression tree (CART), was used to assess the prognostic value of interactions between the standard clinico-pathological variables and the genetic variants in relation to their impact on five-year survival in CRC patients. Multiple testing corrections were performed using the Bonferroni test or the Benjamini-Hochberg false discovery rate.

Detailed information about individual patients, methods of sample processing and methods of individual analyses are reported in the enclosed publications.

4. RESULTS & DISCUSSION

The subject of this Thesis was to investigate the genetic variability in patients with sporadic form of CRC in association with the risk of disease development, as well as with survival prognosis and different treatments response prediction.

The working hypotheses and the experimental work were driven by several major assumptions: (1) Genetics plays a key role in predisposition to CRC, its initiation, and progression. SNPs in candidate genes (DNA repair genes – **Manuscript I and VII**, genes involved in CRC pathogenesis – **Manuscript II**, mucin genes – **Manuscript III** and cancer-related immunity genes – **Manuscript IV, V and VI**) may alter the final protein function and/or efficiency and thus induce genetic instability and unregulated cell growth therefore further influence CRC susceptibility, patient's survival and efficacy of chemotherapy. (2) Genetic variants within miRNA binding sites of targeted genes may cause an altered binding of specific miRNAs to the 3′UTR and thus might be responsible for an aberrant gene expression ultimately affecting CRC risk and modulate the clinical outcome after cancer diagnosis. (3) Understanding the SNPs effect on the CRC risk, survival and treatment response might result in low-cost and low-invasive prognostic and predictive biomarkers with a potential in helping to define individual CRC risk and tailor disease management based on the unique molecular profile of each patient. Individualized therapy would eventually help to improve therapeutic efficacy and minimize toxicities. (4) The association studies are conducted on a considerable number of cases and controls, homogeneous for their ancestry, and clinically well-defined, thus minimizing any possible population stratifications.

In this section, the major findings from each publication representing the PhD study are discussed.

Manuscript I:

The study "*Double-strand break repair and colorectal cancer: gene variants within 3' UTRs and microRNAs binding as modulators of cancer risk and clinical outcome*" explored the association of 21 polymorphisms in 3′UTRs in relevant DNA repair genes (*RAD51, RAD52, BRCA1, MRE11A, NBN, GEN1, XRCC2, XRCC4, XRCC5, LIG4, and NHEJ1)* of DSB repair pathway with CRC susceptibility and prognosis.

The major finding of the study comprising 1126 cases and 1469 controls identifies the carriers of the variant CC genotype in *MRE11A* rs2155209 as strongly associated with a decreased risk of CRC ($p = 0.0004$). Further, a potential SNP-SNP interaction in modulating CRC susceptibility revealed a tendency for the under-representation of cases in comparison with controls among carriers of the variant rs2155209 genotype CC in *MRE11A* in combinations with other SNPs (rs3218547 and rs1051669). In the survival analyses among CRC patients, and specifically among those with colon cancer, carriers of the variant CC genotype in *MRE11A* rs2155209 showed a worse survival when compared with the most frequent TT genotype ($p = 0.03$) (Figure 12). As the SNP rs2155209 in *MRE11A* gene appeared as important in both, the risk and survival analysis, its role in modulating *MRE11A* expression was further investigated by a Dual-Luciferase Reporter Assay. Between the two constructs carrying the different alleles of the SNP a statistically significant difference was observed ($p = 0.007$) and the C-allele was related to a reduced activity of the reporter gene by 14% of the average luciferase activity in comparison with the values obtained for the construct with the T-allele (Figure 13).

Since miRNAs have been recognized as pivotal players in diverse biologic processes (including DNA repair and DNA damage response (Chowdhury et al. 2013; Sharma and Misteli 2013)), the presence of SNPs within the 3′UTRs of target DNA repair genes might cause an altered binding of specific miRNAs, and thus modulate gene expression and ultimately affect cancer susceptibility (Naccarati et al. 2012; Slaby et al. 2012; Cipollini et al. 2014), therapy outcomes (Teo et al. 2012) and survival (Pardini et al. 2013). For example, SNP in miRNA binding site within the DNA repair gene *RAD51* has been reported associated with bladder cancer risk and radiotherapy outcomes (Teo et al. 2012).

MRE11A as part of MRE11-RAD50-NBS1 (MRN) complex is involved in several important processes including a DSB repair (Williams et al. 2010). Mutations in the

complex components have been reported in acute lymphoblastic leukemia, head and neck, prostate, breast and colorectal cancers (Mosor et al. 2006; Dzikiewicz-Krawczyk 2008; Ziolkowska-Suchanek et al. 2013). Several genetic variants in *MRE11* (not in linkage with rs2155209) have also been associated with various cancers including breast, bladder and ovarium (Bartkova et al. 2008; Dzikiewicz-Krawczyk 2008; Chowdhury et al. 2013; Teo et al. 2014). Regarding the studied miRSNP rs2155209 in *MRE11A* gene, the association has been previously reported with an increased risk of myocardial infarction, breast and bladder cancer (Choudhury et al. 2008; Verschuren et al. 2013; Wu et al. 2015).

Protein expression levels of *MRE11* have also been measured in previous studies. An overexpression was commonly observed among CRC patients, therefore it has been postulated as a mechanism responsible for increasing cancer risk (Cancer Genome Atlas 2012). This hypothesis is also supported by RNA sequencing data of CRC patients in The Cancer Genome Atlas (TCGA) database where a general overexpression of all available *MRE11A* transcripts was observed in the tumor tissues when compared with their nonmalignant tissue counterparts (Cancer Genome Atlas 2012).

Since the low-risk allele (C) is associated with a lower expression of *MRE11A* as suggested by the results of our functional study, our outcomes are in accordance with the abovementioned data. Therefore, we may hypothesize that a miRNA post-transcriptional regulation of *MRE11A* may be finely modulated by the presence of the identified miRSNP, with the CC genotype contributing to a reduced risk of developing CRC.

Furthermore, carriers of the *MRE11A* rs2155209 variant CC genotype showed a shorter survival. Notably, MRE11 protein deficiency has been observed to be associated with improved survival of stage III colon cancer patients, independently of treatment (Pavelitz et al. 2014). This study supports our finding where CC genotype of *MRE11A* rs2155209 is associated with shorter survival. We can theorize that the modulatory role by the observed miRSNP on the expression of MRE11 protein may also influence the prognosis of cancer.

In this study, we have reported a significant role of SNP rs2155209 in miRNA target site of DSB repair gene *MRE11A* in CRC risk and clinical outcome, and our results support the idea of miRNAs and miRSNPs contribution to CRC (Vishnubalaji et al. 2015). Since the interest on miRNAs has lately increased for the possibility to use them as diagnostic, prognostic and predictive clinical biomarkers (Iorio and Croce 2012), a

similar study design was applied in **Manuscript II and III** to explore miRSNPs in genes frequently mutated in CRC pathogenesis and in mucin genes.

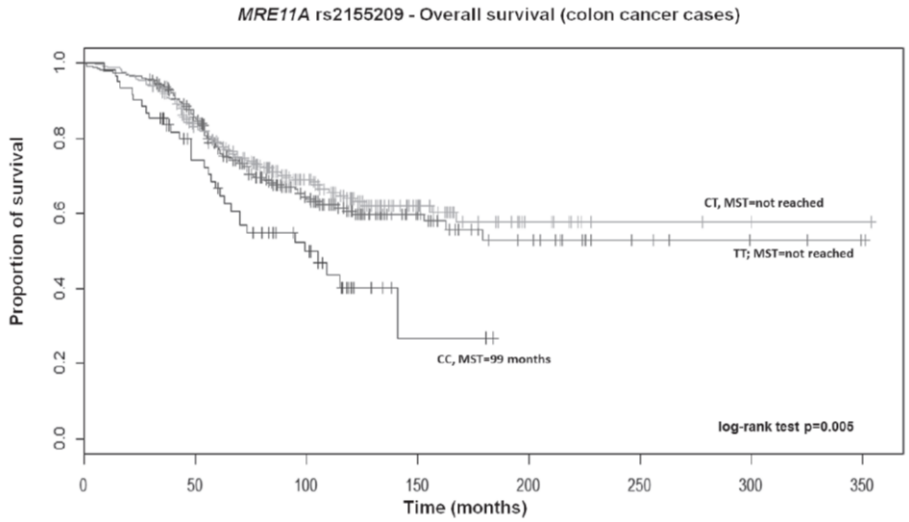


Figure 12: Kaplan-Meier OS curves for *MRE11A* rs2155209 in colon cancer patients.

MST, median survival time.

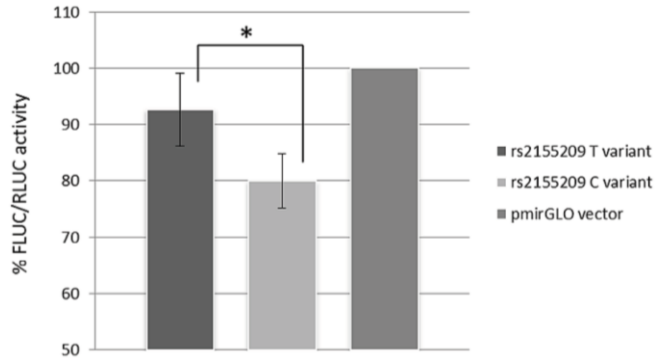


Figure 13: Data show mean values of luminescence activity. *MRE11A* expression shows a statistically significant (p = 0.007) decrease of about 14% in presence of the rs2155209 C-variant, compared to the expression obtained with the T-variant.

Manuscript II:

The study "*MicroRNA-binding site polymorphisms in genes involved in colorectal cancer etiopathogenesis and their impact on disease prognosis*" was performed to test 8 genetic variants in the 3′UTRs of 5 highly penetrant genes (*APC*, *ATM*, *KRAS*, *PARP1*

and *SMAD7*) known to be frequently mutated in CRC pathogenesis in association with CRC risk and clinical outcome.

Selected miRSNPs were tested in the same cohort of patients and controls as described in **Manuscript I**. In a case-control study, the polymorphism rs8679 in *PARP1* gene was associated with a risk of CRC. In particular, the carriers of at least one C allele were at a decreased risk of cancer ($p = 0.05$). The CC genotype in *PARP1* rs8679 was also associated with patient's survival when patients undergoing 5-FU-based chemotherapy were at increased risk of recurrence/progression ($p = 0.03$) (Figure 14).

As already mentioned in the previous **Manuscript I**, the ability of miRNAs to locate and bind a target mRNA has been found to be critical for regulating transcripts level and protein expression necessary for a proper DNA repair and DNA damage response (Preskill and Weidhaas 2013). Thus inherited genetic variants in miRNA target sites are suggested to affect miRNA function and may have an important role in human disease susceptibility and progression (Sethupathy and Collins 2008; Teo et al. 2012; Ryan et al. 2015).

For example, regarding *PARP1* rs8679, a miR-145-3p is predicted to bind to the 3′UTR region where the polymorphism is located. To evaluate whether the studied genetic variant in miRNA target region could potentially alter the binding with specific miRNA, the RNAcofold software was used to predict the difference of binding energy according to the allele (Landi et al. 2012; Naccarati et al. 2012). The result suggests a less efficient binding of miR-145-3p in the presence of the less common C allele and implicates a potentially decreased post-transcriptional repression of *PARP1* by this miRNA. We may therefore hypothesize that a decreased CRC risk is associated with a variant C allele because of a less efficient binding of the miR-145-3p causing an increased level of *PARP1*. This may be supported by a study on CRC cell lines where overexpression of miR-145 was observed as associated with inhibition of cell proliferation, motility and invasion and a stable overexpression of miR-145 suppressed tumor growth and pulmonary metastasis also *in vivo* (Feng et al. 2014). Furthermore, several studies previously reported this miRNAs downregulation of expression in CRC thus adding emphasis to our hypothesis (Feng et al. 2014; Gattolliat et al. 2015; Ramzy et al. 2015).

However, when we utilize a MicroSniper software and focus on miRNAs predicted to bind in the same position in presence of both alleles for *PARP1* rs8679, only one (miR-3074-5p) out of six exhibited the highest binding energy necessary for binding to the

region in the presence of the rare C allele. Thus it might be globally more favorable for miRNAs to bind the rs8679 3′UTR target region when the C allele is present, which might result in a more stringent repression of translation (i.e. decreased target gene expression). Concerning these specific miRSNPs, our findings observing that carriers of at least one C allele are at a decreased risk of CRC are in disagreement with this computational prediction. Nevertheless, it supports once again our initial hypothesis of different allele specificity on miRNA-binding target sites that may be reflected in miRNA regulation.

The CC genotype in rs8679 was further associated with an increased risk of recurrence or progression in patients that received 5-FU-based chemotherapy. We may hypothesize that SNPs within the *PARP1* gene might lead to a decrease in its activity eventually impacting in the failure of apoptosis. Thus the effectiveness of apoptotic activity after 5-FU treatment could be lower, leading to a worse prognosis. This might be supported by a study by Lu *et al.* when increased expression levels of miR-335, predicted *in silico* to bind to C allele of rs8679, were markedly associated with CRC tumor size and differentiation (Lu et al. 2016). In the study by Cheng and colleagues, a close association between *PARP1* gene and 5-FU-based chemotherapy was also described (Cheng et al. 2012).

Understanding the modulating effect of miRNAs on PARP1 protein levels in CRC tumors is particularly important because of the current interest in the use of *PARP1* inhibitors as a single agent or as a chemo- or radiosensitizer in cancer treatment (Megnin-Chanet et al. 2010). In the present study, we provide evidence that variations in potential miRNA-binding target sites in the 3′ UTR of PARP1 gene may modulate CRC risk and prognosis after therapy.

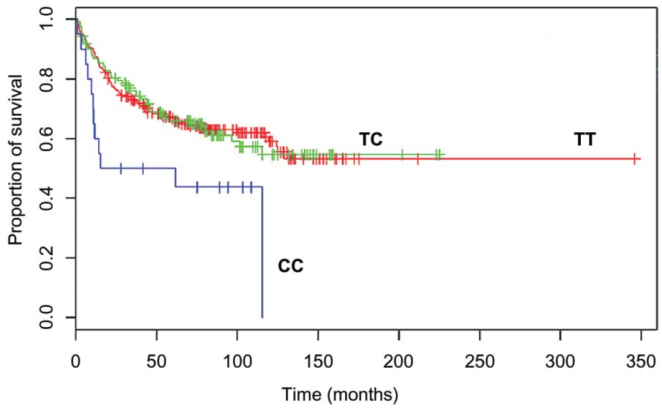


Figure 14: Kaplan–Meier EFS curves in CRC patients undergoing 5-FU-based chemotherapy stratified for rs8679 in *PARP1* gene.

Manuscript III:

The study "*Polymorphisms in microRNA binding sites of mucin genes as predictors of clinical outcome in colorectal cancer patients*" describes the association of 13 miRSNPs of 9 mucin genes (*MUC6*, *MUC7*, *MUC13*, *MUC14*, *MUC15*, *MUC17*, *MUC20, MUC21* and *MUC24*) with CRC risk and clinical outcome.

Selected miRSNPs were assessed in the same cohort of patients and controls as in **Manuscript I and II**. Overall, no strongly significant associations were observed in the case-control study. Borderline significant p-values for decreased risk of CRC were found for 4 miRSNPs (*MUC13* rs1532602, *MUC14* rs4071, *EMCN/MUC14* rs17552409 and *MUC24* rs974034) either in the entire group of CRC patients or after stratification according to the tumor site. Among the strongest associations with patient's survival, the carriers of the CC genotype in *MUC21* rs886403 were associated with a worse survival and a higher recurrence risk in CRC patients (OS: $p = 0.01$ and EFS: $p = 0.0002$), which was even more pronounced in colon cancer cases (OS: $p < 0.0001$ and EFS: $p < 0.0001$) (Figure 15A). In contrast, rectal cancer patients carrying the variant CC genotype in *MUC17* rs4729655 displayed a better survival ($p = 0.0002$) (Figure 15B). Finally, CRC and colon cancer patients carrying the variant GG genotype of in *MUC20* rs6782006 showed a worse OS ($p = 0.02$).

Mucins are glycoproteins predominantly expressed at the epithelial part of tissues and provide a protection for colon surface. Under physiological conditions, mucins maintain a homeostasis by covering human colon surface by gel mucous layer (Gupta et al. 2012). During the malignant development, miRNAs have emerged as important regulators responsible for an altered mucin expression (Macha et al. 2015). Therefore, miRSNPs located in mucin genes may play a role in cancer susceptibility, efficacy of chemotherapy and survival.

Several mucin genes, such as *MUC21*, *MUC17*, and *MUC20*, appeared to be significantly associated with patient's survival in our study. Unfortunately only scarce information is available for mucin genes in the published articles. Concerning *MUC21*, the study by Yi *et al.* showed significantly less adherent cells to each other and to extracellular matrix components in cells transfected by *MUC21* (Yi et al. 2010), suggesting a role of *MUC21* in cell adhesion. The appropriate cell adhesion is necessary for numerous physiological processes (cellular organization, differentiation, proliferation and survival). It also plays vital roles in many later steps in cancer

progression (entry of cancer cells into the bloodstream and their establishment at distant organs) (Labelle and Hynes 2012). Another mucin *MUC17* is normally highly expressed on the surface epithelium of colon/rectal mucosa but its expression becomes altered in colorectal neoplasia. Li *et al.* observed an increased expression of *MUC17* associated with a longer OS in patients with stage III and IV of colorectal adenocarcinomas (Li 2011) which points to its possible role in cancer progression and prognosis. In the study by Kitamoto *et al.* several miRNAs were proposed as potential regulators of *MUC17* expression, but none of them has been validated *in vivo* (Kitamoto et al. 2011). Lastly, *MUC20* gene is highly expressed in kidney and colon tissues. Xiao *et al.* observed a significantly upregulated *MUC20* in CRC patients with poor prognosis (Xiao et al. 2013) and a relationship between overexpression and poor survival was also found in other human malignancies (ovarian cancer, non-small cell lung cancer, endometrial cancer and gastric cancer) (Vlad et al. 2006; Woenckhaus et al. 2008; Marin et al. 2012; Chen et al. 2013).

In the present study, plausible candidate miRSNPs potentially affecting miRNA binding to mucin genes were identified as related to either CRC susceptibility or patient´s survival. Our results support the idea that a 'miRNA network' may contribute to CRC pathogenesis. Expanding our knowledge on mucins may help us to better understand the etiopathogenesis of CRC and thereby contribute to the development of new treatment strategies.

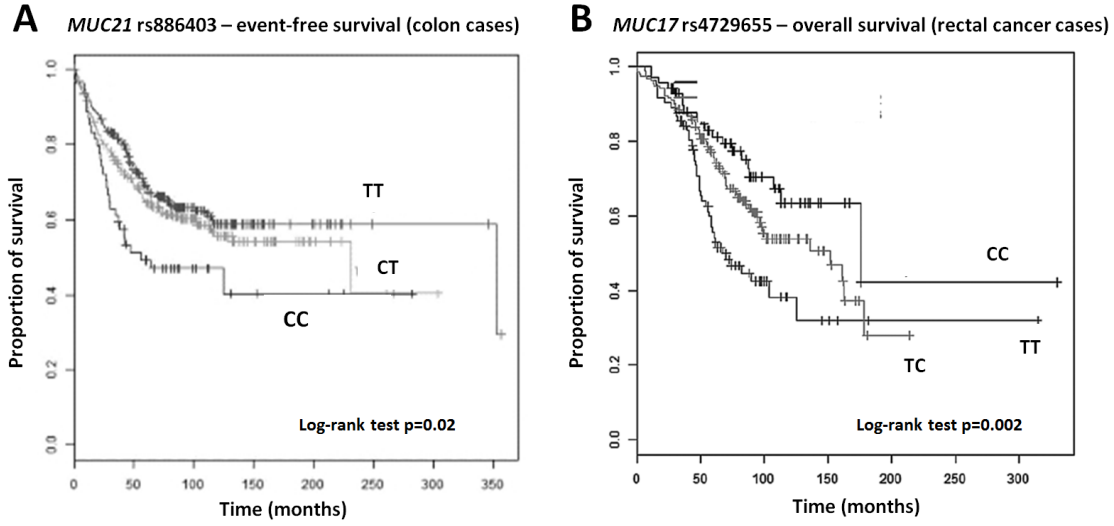


Figure 15: Kaplan–Meier curves **(A)** EFS for rs886403 in *MUC21* gene in colon cancer patients **(B)** OS for rs4729655 in MUC17 gene in rectal cancer patients.

Manuscript IV:

The study "*Coding variants in NOD-like receptors: An association study on risk and survival of colorectal cancer*" was aimed at evaluating 41 non-synonymous SNPs in 21 NLR genes (*NLRP1-14, NLRC4* and *5, NOD1* and *2, NAIP, RIPK2* and *ASC*) for their association with CRC risk and clinical outcome.

In this study, a discovery cohort from the Czech Republic (1237 cases and 787 controls) and two large GWAS data sets for a replication analysis were included (Germany: 1798 cases and 1810 controls and Scotland: 2210 cases and 9350 controls). The major findings in a discovery set describe five SNPs to be significantly associated with CRC risk (rs1043673, rs35829419, rs6421985, rs306457 and rs303997) and eight with patient´s survival (rs12150220, rs1043673, rs10409555, rs12462795, rs16986899, rs34436714, rs289723 and rs74439742), however the associations were not confirmed in a replication analysis. To assess the expression of NLRs found in the Czech discovery set in the gut or immune cells, mRNA levels were measured in primary tissue samples and cell lines and divergent expression patterns of *NLRP2*, *5*, *6* and *13* were found in hematopoietic and non-hematopoietic cells (Figure 16).

NLRs are important innate pattern recognition receptors and regulators of inflammation and play an important part in the homeostasis of the immune system (Wilmanski et al. 2008; Oviedo-Boyso et al. 2014). A different expression of NLR genes may lead to a disruption of the underlying regulatory pathways and result in the development of local and chronic inflammation, inflammatory bowel disease and/or CRC (Abreu 2010; Carvalho et al. 2012).

In spite of the *in silico* predictions about the functionality of studied SNPs the promising results from the Czech cohort could not be confirmed in the two GWAS data sets. However, our expression analysis showed the NLRs associated with CRC risk or survival in the discovery set as expressed in primary human colon or rectum cells, CRC tissue and cell lines, providing preliminary evidence for a potential involvement of NLRs in CRC development and progression. Furthermore, the expression of development-related *NLRP5* was undetectable in nonmalignant colon tissue but was upregulated in colon cancer tissue and cell lines, suggesting a potential novel role beyond developmental control for this NLR in humans (Lupfer and Kanneganti 2013). Induced expression of *NLRC5* in HCT116 cells may have a functional outcome by modulating major histocompatibility complex (MHC) class I expression (Neerincx et al.

2013) which correlates with survival due to its effect on CD8 cytotoxic T-cell and natural killer cell immuno-surveillance (Watson et al. 2006).

To further uncover the poorly understood role of NLRs in CRC development and survival, the effect of regulatory variants of *NLRC5* on CRC susceptibility and clinical outcome were explored in **Manuscript V and VI**.

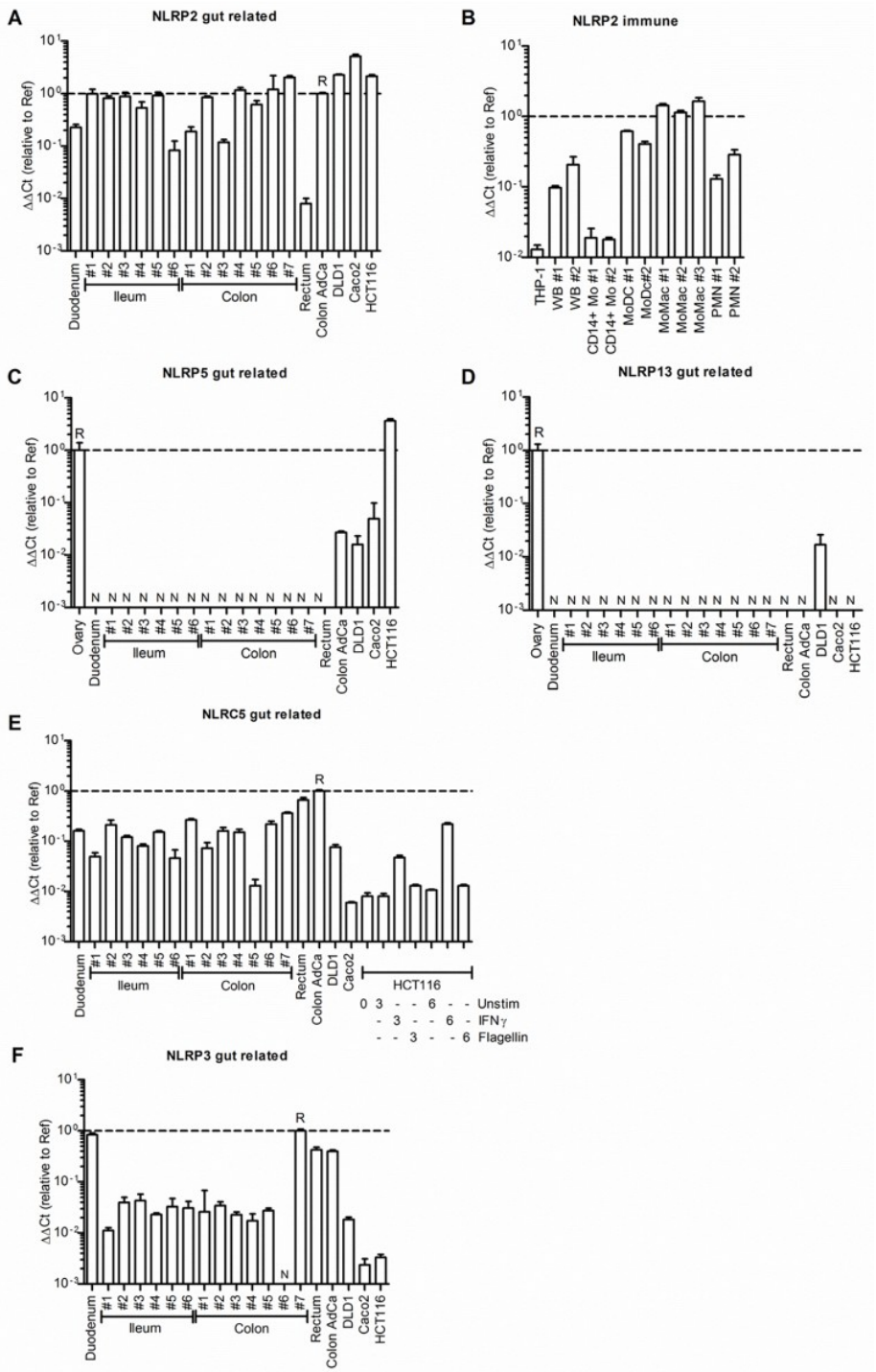


Figure 16: Expression of selected CRC-associated NLRs in immune cells, primary tissue samples or CRC cell lines.

Manuscript V:

The study "*Investigation of single and synergic effects of NLRC5 and PD-L1 variants on the risk of colorectal cancer*" reports the influence of 16 potential regulatory variants in the *NLRC5* and Programmed death-ligand 1 (*PD-L1)* genes selected by several *in silico* tools on CRC susceptibility.

This case-control study comprising 1424 cases and 1114 healthy controls from the Czech Republic reports a moderate association between rectal cancer risk and two *NLRC5* SNPs (rs1684575 and rs3751710). Given that the evaluated SNPs did not show any strong individual association with CRC risk and that a combination of genetic variants, rather than a single polymorphism, may explain better the genetic etiology of CRC, we further focused on the interplay between the variants. Eighteen pair-wise interactions within and between the *NLRC5* ad *PD-L1* genes were obtained. Six more interactions appeared when the previously genotyped *IFNGR1* and *IFNGR2* variants were added to the analysis (Lu et al. 2014). The main interactions included three *NLRC5* SNPs (rs289747, rs289748, rs56315364) located in the upstream and promoter region with the same *PD-L1* promoter SNP rs2890657 (Figure 17).

NLRC5 gene, a member of a NLR family, plays a prominent role in antitumor immunity while *PD-L1* acts as a physiological feedback mechanism necessary for terminating the immune responses and for maintaining self-tolerance (Riella et al. 2012). Changes in these genes expression may lead to a disrupted anti-tumor immune response, which in turn may influence CRC susceptibility (Lynch and Murphy 2016; Yoshihama et al. 2016; Passardi et al. 2017). SNPs located in regulatory regions of *NLRC5* and *PD-L1* may thus be actively involved in the regulation of gene expression and have an impact on CRC development (Khurana et al. 2016).

As *NLRC5* has been reported to be a MHC class I transactivator, its upregulated expression could lead to a strong CD8+ activation necessary for generating an effective immune defense against invading harmful pathogens (Pandiyan et al. 2007). On the other hand, a downregulated expression of *NLRC5* has been reported to lead to an impaired ability to elicit CD8+ T-cell activation, which represents a way used by the tumor cells to escape the host immune system (Yoshihama et al. 2016). Promoted *PD-L1* expression may however negatively regulate primed CD8+ T cell expansion therefore an aberrant *PD-L1* expression might allow cancer cells to escape the antitumor immune response by suppressing the CD8+ T cell expansion (Karwacz et al. 2011;

Riella et al. 2012; Kataoka et al. 2016). Additionally, the data from the study by Van *et al.* suggest that 5-FU, a chemotherapeutic frequently used in CRC treatment, has an impact on *PD-L1* expression (Van Der Kraak et al. 2016).

Our data suggest that the interaction between the inherited genetic variants within genes involved in immune surveillance contributes to signaling defects, which in turn may lead to alteration in the anti-tumor immune response and further play an important role in the onset of CRC. Expanding our knowledge on regulatory variants in the *NLRC5* and *PD-L1* genes could eventually improve CRC risk management but also *PD-L1*-based immunotherapy in CRC. The association of *NLRC5* with regard to a therapy response as well as survival of the CRC patients is further explored in **Manuscript VI.**

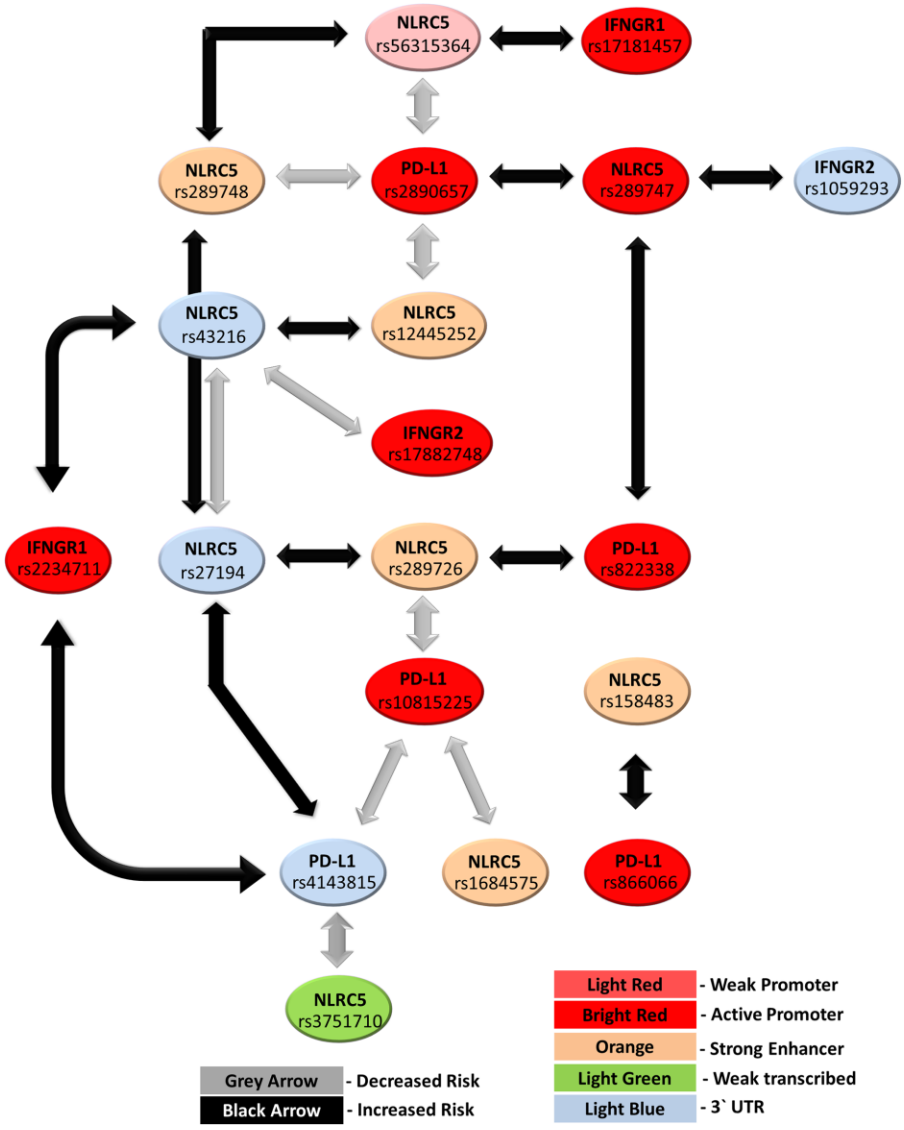


Figure 17: *NLRC5-PD-L1-IFNGR1/2* pair-wise interactions. The color indicates the SNPs' location displayed by UCSC Genome Browser on lymphoblastoid cell lines (GM12878).

Manuscript VI:

The study "*Influence of regulatory NLRC5 variants on colorectal cancer survival and 5-fluorouracil-based chemotherapy*" was preformed to evaluate the effect of 11 potential regulatory polymorphisms in *NLRC5* selected by several *in silico* tools on overall and event-free survival of patients with sporadic form of CRC.

The case-only study was carried out on 589 CRC cases from the Czech Republic (232 patients received 5-FU-based therapy). Minor alleles of two SNPs (rs27194 and rs289747) were significantly associated with a decreased survival in all patients and metastasis-free patients at the time of diagnosis (pM0). Among CRC patients receiving a 5-FU-based adjuvant therapy, rs12445252 was associated according to the dosage of the minor allele T with OS and EFS (Figure 18).

As previously mentioned *NLRC5* is a transactivator of MHC class I molecules (Meissner et al. 2010) and plays a pivotal role in immune-surveillance with a potential influence on cancer patient's survival. It is an interferon gamma (IFNγ)–inducible nuclear protein and due to the link with the IFNγ system, *NLRC5* might also play a role in the 5-FU-based therapy.

The SNP rs27194, associated with a decreased survival in our study, is located in a 3'UTR region and therefore by affecting miRNA binding sites might be responsible for aberrant *NLRC5* expression in CRC. This hypothesis may be supported by results of Microsniper software, when miR-942 (predicted as possibly affected by miRSNP rs27194) is known to be constitutively activated in many cancers, including CRC (Zhan et al. 2017). Furthermore, the same miRSNP is predicted to affect the binding site affinity of *PRRX2* and *TCF4*, both strongly deregulated in CRC (Xu and Pasche 2007; Zhan et al. 2017). The second SNP rs289747 associated with the survival is an intronic variant which is presumed to affect an OCT1 binding site by increasing its binding affinity. Also *OCT1* overexpression has been described in many cancers, including CRC (Wang et al. 2016). Lastly the rs12445252 polymorphism, found to be associated with 5-FU treatment survival, is an intronic eQTL variant negatively influencing the expression of *NLRC5* and a decreased expression of *NLRC5* could further affect chemotherapeutic efficacy of 5-FU.

Unfortunately, studies on expression levels of *NLRC5* in CRC and normal tissue are sparse and contradictory (Liu et al. 2015; Yoshihama et al. 2016). Only a few studies have addressed the role of *NLRC5* gene expression and survival of cancer patients while

results have indicated high expression of *NLRC5* as a good prognostic marker (including CRC) (Watson et al. 2006; Simpson et al. 2010).

Our results indicate that polymorphisms in immune surveillance genes, such as *NLRC5*, may be used as prognostic markers for clinical outcome in CRC, as well as for survival of CRC patients in response to 5-FU-treatment. Our study also adds a new layer on the complex function of *NLRC5* in the innate immune system.

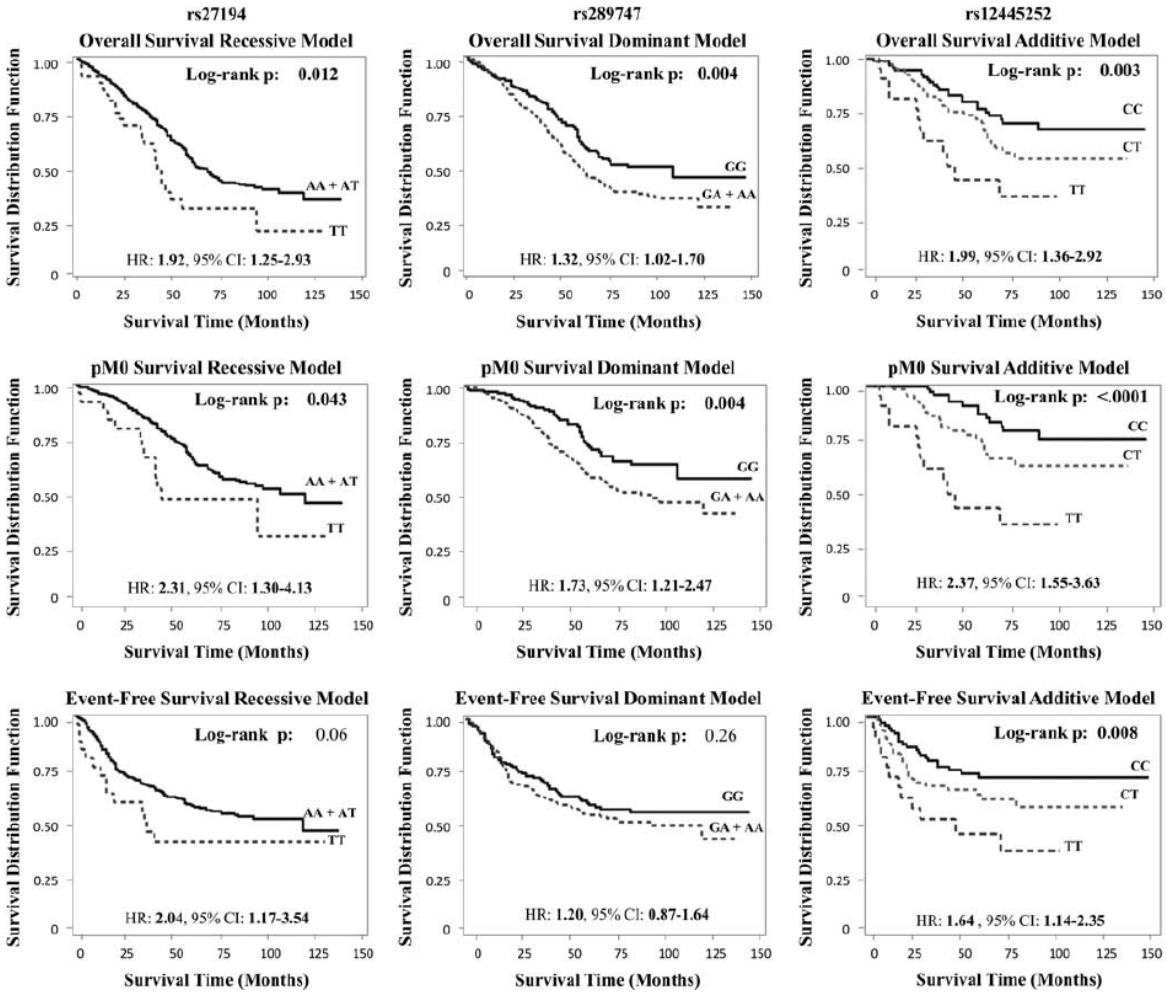


Figure 18: Kaplan–Meier analysis of survival according to genotypes of rs27194 and rs289747 in the whole study population and rs12445252 in 5-fluorouracil-treated.

HR, Hazard ratio; CI, confidence interval

Manuscript VII:

The study "*Functional polymorphisms in DNA repair genes are associated with sporadic colorectal cancer susceptibility and clinical outcome*" evaluated the relevance of 16 functional genetic variants in 12 DNA repair genes (*EME1, FAAP24, FANCI, MUS81, NEIL3, POLE, POLN, POLQ, RAD51D, REV1, REV3L* and *RPA1*) on the risk of CRC development and modulation of the clinical outcome after cancer diagnosis.

In the discovery set of 1832 patients and 1172 controls from the Czech Republic, the carriers of the variant *AA* genotype in *REV3L* rs3204953 (Val2986Ile) were observed as associated with an increased risk of CRC ($p = 0.006$). The valine to isoleucine substitution has been recognized via *in silico* approach, performed by F-SNP database, to have a high probability of being functionally significant (Lee and Shatkay 2008) and in the area of a molecular epidemiology, there is evidence that polymorphisms in *REV3L* are associated with different malignancies. For example, an association of rs3204953 was observed with a higher risk of breast cancer in a Swedish cohort (Varadi et al. 2011) and other genetic variants in *REV3L* have been found to be associated with breast cancer, stomach cancer, and CRC (Hussain et al. 2009; Varadi et al. 2011; Pan et al. 2012). In addition to the prediction of the deleterious nature of the protein function, the amino acid change *REV3L* Val2986Ile was predicted to decrease the protein stability by a web-server ELASPIC (Witvliet et al. 2016). The importance of the accurate level of the functional protein in cells was demonstrated on disrupted REV3L in cancer cell lines, when its inhibition induced a growth arrest in cancer cells, whereas overexpression led to increased spontaneous mutation rates (Knobel and Marti 2011). A decreased expression levels have also been reported in tumor compared with the adjacent nonmalignant tissue in colon cancer (Brondello et al. 2008; Stallons and McGregor 2010). Unfortunately, despite the promising results in the Czech population, an association of *REV3L* SNP with CRC risk could not be confirmed in the Austrian replication set comprising 950 patients and 820 controls. However, since *REV3L* was observed as significant in the Austrian survival analyses and according to all of the available data, we suggest that the *REV3L* gene may impact CRC susceptibility, survival, and therapy outcomes and warrants further investigation.

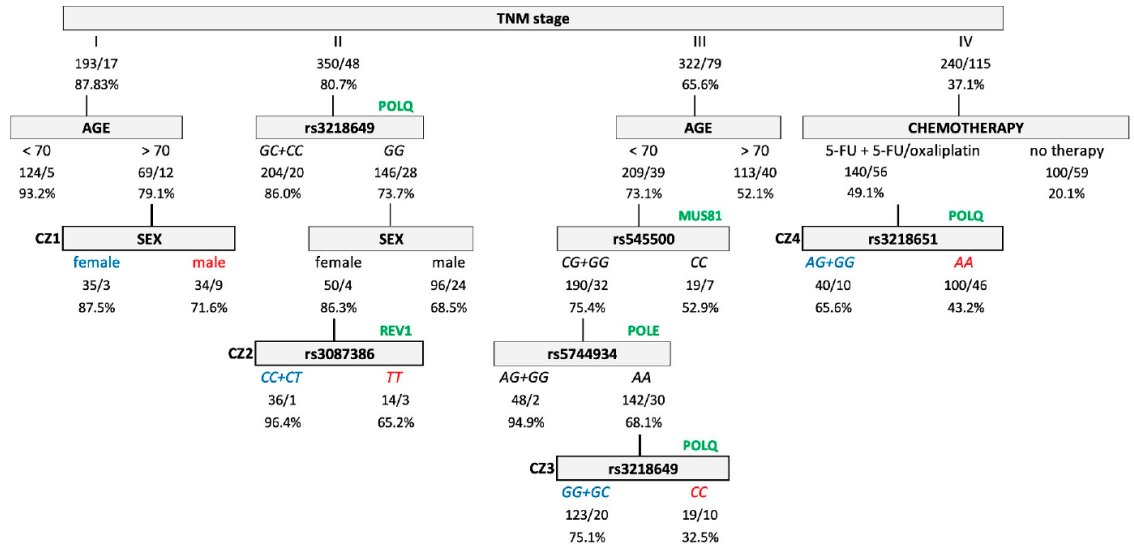
The CART analysis, investigating the interactive effects of genotypes and clinico-pathological parameters in association with five-year OS and EFS, showed a prognostic utility of several investigated DNA repair gene polymorphisms. Only a few of these

were shown as significant more than once in the final structure of the tree, suggesting their potentially greater relevance on patient's survival. *POLQ* gene polymorphisms appeared four times as the optimal split factor in the Czech CART analyses (rs1381057, rs3218649 twice, and rs3218651) and four times in the Austrian CART analyses (rs1381057 twice and rs3218651 twice) (Figure 19). At least nine out of 23 known *POLQ* gene polymorphisms in the human are predicted to alter protein function (Beagan and McVey 2016) and several SNPs have also been associated with a risk of different tumors (breast cancer, esophageal cancer, and Non-Hodgkin`s Lymphoma) (Varadi et al. 2011; Li et al. 2013; Brandalize et al. 2014; Rendleman et al. 2014; Family et al. 2015). Apart from the deleterious nature of the protein function of all studied *POLQ* SNPs predicted by F-SNP database, ELASPIC estimated the substitution of glutamine to arginine of *POLQ* SNP rs1381057 to decrease the final protein stability. The accurate level of the functional protein in cells has been reported as important in a complementary body of literature. An upregulation of *POLQ* was found in different tumor tissues (breast cancer, non-small cell lung cancer, oral squamous cell carcinoma, stomach cancer, and CRC), and this overexpression was associated with the disease prognosis (Kawamura et al. 2004; Lemee et al. 2010; Pillaire et al. 2010; Allera-Moreau et al. 2012). Based on the information from published studies we consider the significance of adequate *POLQ* functioning and regulation for tumor suppression.

In the five-year EFS CART analysis *NEIL3* gene SNP rs7689099 emerged twice as the optimal split factor in the Czech cohort (Figure 20). Different *NEIL3* gene polymorphisms were associated with the risk of several malignancies such as glioma, prostate, and thyroid cancer (Bethke et al. 2008; Barry et al. 2011; Cipollini et al. 2016), with rs7689099 being associated with a reduced risk of differentiated thyroid carcinoma and prostate cancer (Barry et al. 2011; Cipollini et al. 2016). Likewise in previously mentioned *REV3L* and *POLQ*, significantly elevated expression levels of *NEIL3* were reported in tumors of 20 cancer sites, including CRC (Hildrestrand et al. 2009; Shinmura et al. 2016). In case of melanoma, the overexpression was further observed in association with the progression to distant metastasis (Kauffmann et al. 2008). The association of *NEIL3* SNP with the survival of CRC patients was not replicated in the Austrian sample set. However, considering the available data, we suggest that the variation of the *NEIL3* gene also has relevance for CRC susceptibility, survival, and therapy outcome.

In this study, we evaluated the association of genetic variants in DNA repair genes selected by likely functional relevance with CRC. Our data suggest that even subtle alterations in the specific proteins that function in DNA repair pathways may lead to inaccurate DNA repair and thus play a role in CRC pathogenesis.

(A)



(B)

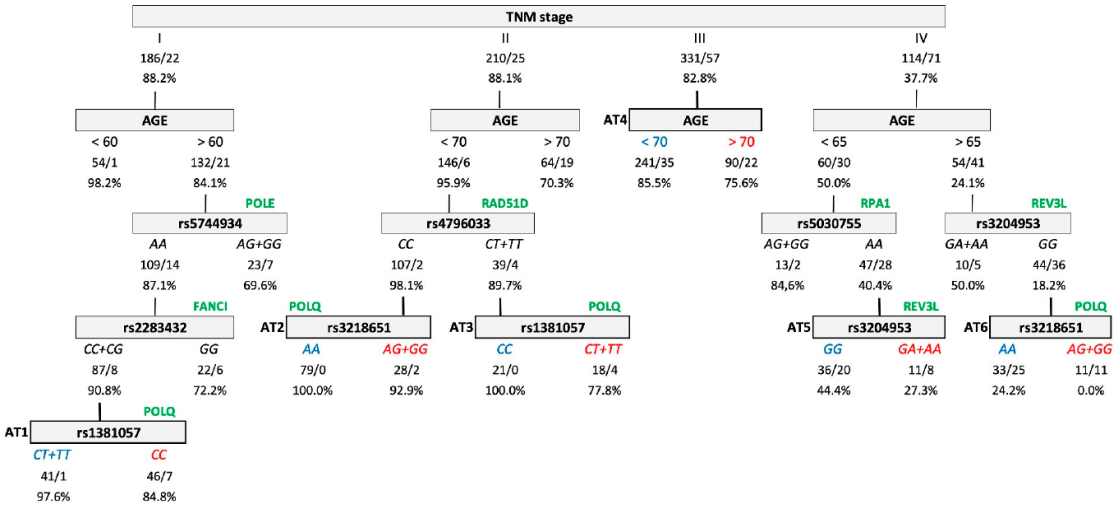
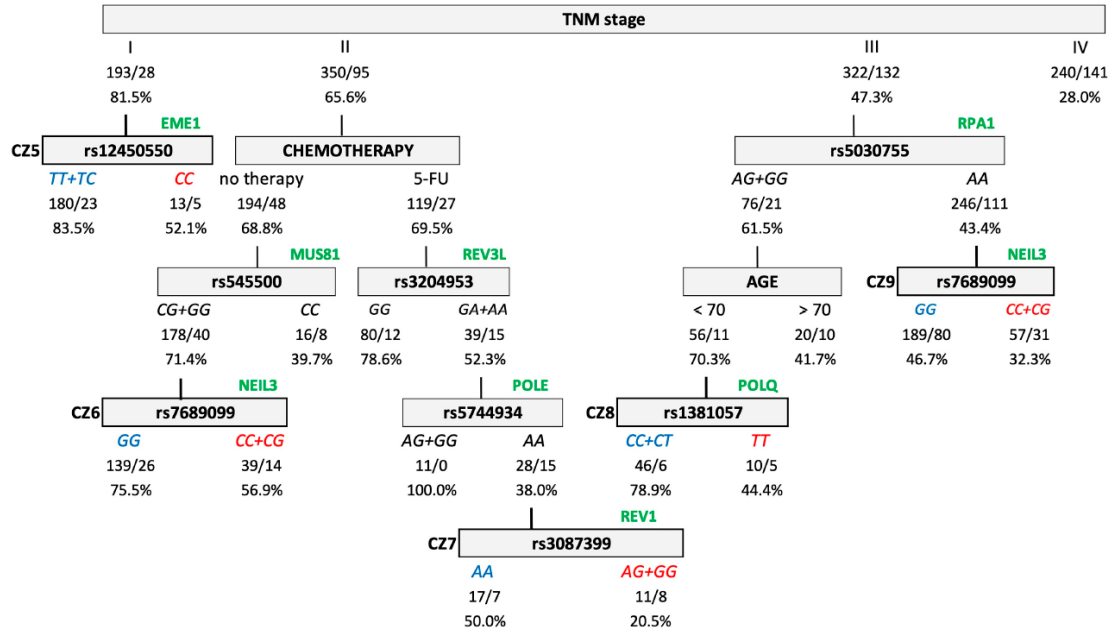


Figure 19: OS classification and regression tree analysis of colorectal cancer patients from the Czech Republic **(A)** and Austria **(B)**. Numbers under each node indicate the total number of cases in the subcategory/number of events and percentage of patients with five-year OS.

(A)



(B)

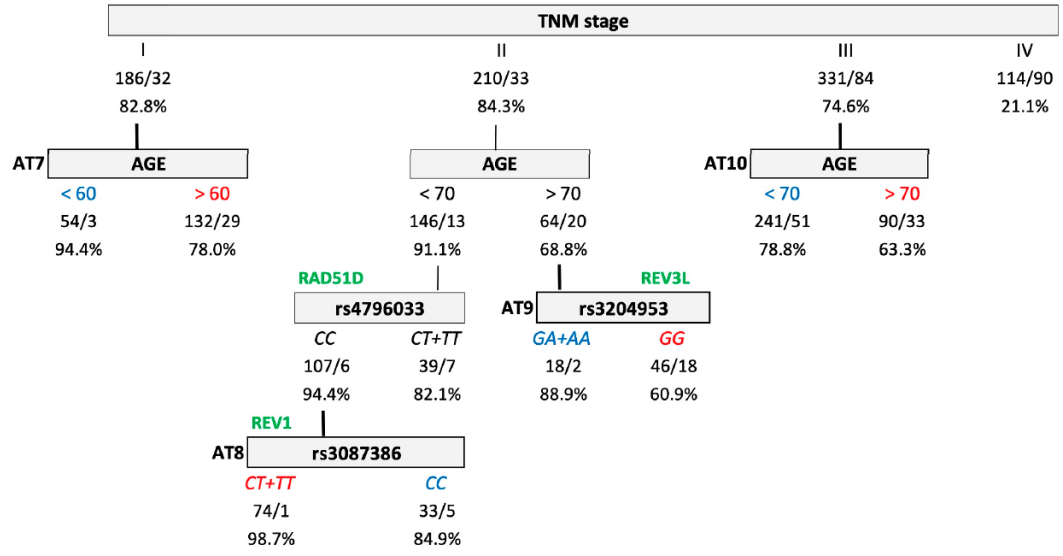


Figure 20: EFS classification and regression tree analysis of colorectal cancer patients from the Czech Republic **(A)** and Austria **(B)**. Numbers under each node indicate the total number of cases in the subcategory/number of events and percentage of patients with five-year EFS.

5. CONCLUSIONS

The main results which were obtained during the work on this thesis are summarized in the following paragraphs:

- We have observed several candidate SNPs in miRNA target regions related either to CRC risk or to clinical outcome. In DSB repair genes, we identified a miRSNP *MRE11A* rs2155209 as strongly associated with a decreased risk of CRC and worse survival (Manucript I). In genes important for CRC etiology, an association between *PARP1* rs8679 and either a decreased risk of CRC or an increased risk of recurrence or progression in patients that received 5-FU-based chemotherapy was observed (Manuscript II). In mucin genes, *MUC21* rs886403 was associated with a worse survival and a higher recurrence risk in CRC patients, *MUC20* rs6782006 showed a worse OS in CRC patients and *MUC17* rs4729655 displayed a better OS in rectal cancer patients (Manuscript III).

- We have found that genetic variations in the 3′ UTR of target genes modulate the efficiency of corresponding protein expressions. We investigated a role of miRSNP *MRE11A* rs2155209 in modulating *MRE11A* expression by a Dual-Luciferase Reporter Assay and a statistically significant difference was observed between two constructs carrying the different alleles of the SNP. One allele was related to a reduced activity of the reporter gene by 14% (Manuscript I).

- We have demonstrated that genetic variants in NLR genes contribute to CRC onset and progression of the disease. In Manuscript IV, 5 SNPs were described to be associated with CRC risk, and eight with CRC survival in the Czech population. Also an additive effect on CRC risk and survival was detected, resulting in a 2-fold increased risk and a 3-fold worse survival for carriers of 6 and 8 risk alleles, respectively. However, the results could not be confirmed in the German and Scottish GWAS data sets and future studies are needed to validate the results.
  In manuscript V and VI, role of SNPs within *NLRC5* gene in CRC risk and clinical outcome was reported. The results suggest that genetic variants in immune

surveillance genes, such as *NLRC5*, may serve as candidate prognostic and predictive markers of CRC.

- We have identified the association of several potential functional SNPs in DNA repair genes with CRC. *REV3L* rs3204953 was observed to be associated with an increased risk of CRC and several other SNPs were shown to be associated with OS and EFS in the CART analyses (Manuscript VII). Our data suggest that even subtle alterations of the final protein caused by amino acid substitution may lead to inaccurate DNA repair, and thus contribute to carcinogenesis.

This thesis suggested several potential candidate biomarkers for clinical use. However, further studies are needed to replicate our findings and assess the SNPs in independent populations, to functionally characterize the significant genetic variants and to find the biologic mechanisms underlying the associations.

6. PUBLICATION ACTIVITY

Papers not related to the PhD thesis, published in journals with IF:

1. Kunzmann AT, Proença MA, Vodicka P, Jiraskova K, Schneiderova M, Levy M, Liska V, Buchler T, Vodickova L, Vymetalkova V, Jordao H, McKenna G, Silva AE, Coleman HG and Hughes DJ. Fusobacterium nucleatum tumor DNA levels are associated with survival in colorectal cancer patients. *European Journal of Clinical Microbiology & Infectious Diseases,* 2019 Jul 31. doi: 10.1007/s10096-019-03649-1. **IF 2.591**

2. Kroupa M, Rachakonda SK, Liska V, Srinivas N, Urbanova M, Jiraskova K, Schneiderova M, Vycital O, Vymetalkova V, Vodickova L, Kumar R, and Vodicka P. Relationship of telomere length in colorectal cancer patients with cancer phenotype and patient prognosis. *British Journal of Cancer,* 2019 Aug;121(4):344-350. **IF 5.922**

3. Vodenkova S, Jiraskova K, Urbanova M, Kroupa M, Slyskova J, Schneiderova M, Levy M, Buchler T, Liska V, Vodickova L, Vymetalkova V, Collins A, Opattova A, Vodicka P. Base excision repair capacity as a determinant of prognosis and therapy response in colon cancer patients. *DNA Repair (Amst).* 2018 Dec;72:77-85. **IF 4.461**

4. Kral J, Korenkova V, Novosadova V, Langerova L, Schneiderova M, Liska V, Levy M, Veskrnova V, Spicak J, Jiraskova K, Opattova A, Vymetalkova V, Vodicka P, Jana Slyskova J. Expression profile of miR-17/92 cluster is predictive of treatment response in rectal cancer. *Carcinogenesis*. 2018 Dec 13;39(11):1359-1367. **IF 5.105**

5. Kroupa M, Polivkova Z, Rachakonda S, Schneiderova M, Vodenkova S, Buchler T, Jiraskova K, Urbanova M, Vodickova L, Hemminki K, Kumar R, Vodicka P. Bleomycin-induced chromosomal damage and shortening of telomeres in peripheral blood lymphocytes of incident cancer patients. *Genes Chromosomes Cancer*. 2018 Feb;57(2):61-69. **IF 3.696**

6. Carrai M, Campa D, Vodicka P, Flamini R, Martelli I, Slyskova J, Jiraskova K, Rejhova A, Vodenkova S, Canzian F, Bertelli A, Dalla Vedova A, Bavaresco L, Vodickova L, Barale R. Association between taste receptor (TAS) genes and the perception of wine characteristics. *Sci Rep.* 2017 Aug 23;7(1):9239. **IF 4.259**

7. Vodicka P, Musak L, Frank C, Kazimirova A, Vymetalkova V, Barancokova M, Smolkova B, Dzupinkova Z, Jiraskova K, Vodenkova S, Kroupa M, Osina O, Naccarati A, Palitti F, Försti A, Dusinska M, Vodickova L, Hemminki K. Interactions of DNA repair gene variants modulate chromosomal aberrations in healthy subjects. *Carcinogenesis*. 2015 Nov;36(11):1299-306. **IF 5.334**

8. Vymetalkova V, Soucek P, Kunicka T, Jiraskova K, Brynychova V, Pardini B, Novosadova V, Polivkova Z, Kubackova K, Kozevnikovova R, Ambrus M, Vodickova L, Naccarati A, Vodicka P. Genotype and Haplotype Analyses of TP53 Gene in Breast Cancer Patients: Association with Risk and Clinical Outcomes. *PLoS One.* 2015 Jul 30;10(7):e0134463. **IF 3.234**

9. Vodenkova S, Polivkova Z, Musak L, Smerhovsky Z, Zoubkova H, Sytarova S, Kavcova E, Halasova E, Vodickova L, Jiraskova K, Svoboda M,Ambrus M, Hemminki K, Vodicka P. Structural chromosomal aberrations as potential risk markers in incident cancer patients. *Mutagenesis.*2015 Jul;30(4):557-63. **IF 3.497**

7. REFERENCES

Abba ML, Patil N, Leupold JH, Moniuszko M, Utikal J, Niklinski J, Allgayer H. 2017. MicroRNAs as novel targets and tools in cancer therapy. Cancer Lett 387:84-94.

Abreu MT. 2010. Toll-like receptor signalling in the intestinal epithelium: how bacterial recognition shapes intestinal function. Nat Rev Immunol 10(2):131-144.

Abuli A, Bessa X, Gonzalez JR, Ruiz-Ponte C, Caceres A, Munoz J, Gonzalo V, Balaguer F, Fernandez-Rozadilla C, Gonzalez D, de Castro L, Clofent J, Bujanda L, Cubiella J, Rene JM, Morillas JD, Lanas A, Rigau J, Garcia AM, Latorre M, Salo J, Fernandez Banares F, Arguello L, Pena E, Vilella A, Riestra S, Carreno R, Paya A, Alenda C, Xicola RM, Doyle BJ, Jover R, Llor X, Carracedo A, Castells A, Castellvi-Bel S, Andreu M, Gastrointestinal Oncology Group of the Spanish Gastroenterological A. 2010. Susceptibility genetic variants associated with colorectal cancer risk correlate with cancer phenotype. Gastroenterology 139(3):788-796, 796 e781-786.

Allera-Moreau C, Rouquette I, Lepage B, Oumouhou N, Walschaerts M, Leconte E, Schilling V, Gordien K, Brouchet L, Delisle MB, Mazieres J, Hoffmann JS, Pasero P, Cazaux C. 2012. DNA replication stress response involving PLK1, CDC6, POLQ, RAD51 and CLASPIN upregulation prognoses the outcome of early/mid-stage non-small cell lung cancer patients. Oncogenesis 1:e30.

Amos W, Driscoll E, Hoffman JI. 2011. Candidate genes versus genome-wide associations: which are better for detecting genetic susceptibility to infectious disease? Proc Biol Sci 278(1709):1183-1188.

Andre T, Boni C, Mounedji-Boudiaf L, Navarro M, Tabernero J, Hickish T, Topham C, Zaninelli M, Clingan P, Bridgewater J, Tabah-Fisch I, de Gramont A, Multicenter International Study of Oxaliplatin/5-Fluorouracil/Leucovorin in the Adjuvant Treatment of Colon Cancer I. 2004. Oxaliplatin, fluorouracil, and leucovorin as adjuvant treatment for colon cancer. N Engl J Med 350(23):2343-2351.

Andre T, Boni C, Navarro M, Tabernero J, Hickish T, Topham C, Bonetti A, Clingan P, Bridgewater J, Rivera F, de Gramont A. 2009. Improved overall survival with oxaliplatin, fluorouracil, and leucovorin as adjuvant treatment in stage II or III colon cancer in the MOSAIC trial. J Clin Oncol 27(19):3109-3116.

Andrianifahanana M, Moniaux N, Batra SK. 2006. Regulation of mucin expression: mechanistic aspects and implications for cancer and inflammatory diseases. Biochim Biophys Acta 1765(2):189-222.

Aran V, Victorino AP, Thuler LC, Ferreira CG. 2016. Colorectal Cancer: Epidemiology, Disease Mechanisms and Interventions to Reduce Onset and Mortality. Clin Colorectal Cancer 15(3):195-203.

Arnold M, Sierra MS, Laversanne M, Soerjomataram I, Jemal A, Bray F. 2017. Global patterns and trends in colorectal cancer incidence and mortality. Gut 66(4):683-691.

Barrett JC, Clayton DG, Concannon P, Akolkar B, Cooper JD, Erlich HA, Julier C, Morahan G, Nerup J, Nierras C, Plagnol V, Pociot F, Schuilenburg H, Smyth DJ, Stevens H, Todd JA, Walker NM, Rich SS, Type 1 Diabetes Genetics C. 2009. Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. Nat Genet 41(6):703-707.

Barry KH, Koutros S, Berndt SI, Andreotti G, Hoppin JA, Sandler DP, Burdette LA, Yeager M, Freeman LE, Lubin JH, Ma X, Zheng T, Alavanja MC. 2011. Genetic variation in base excision repair pathway genes, pesticide exposure, and prostate cancer risk. Environ Health Perspect 119(12):1726-1732.

Bartkova J, Tommiska J, Oplustilova L, Aaltonen K, Tamminen A, Heikkinen T, Mistrik M, Aittomaki K, Blomqvist C, Heikkila P, Lukas J, Nevanlinna H, Bartek J. 2008. Aberrations of the MRE11-RAD50-NBS1 DNA damage sensor complex in human breast cancer: MRE11 as a candidate familial cancer-predisposing gene. Mol Oncol 2(4):296-316.

Baxter NN, Goldwasser MA, Paszat LF, Saskin R, Urbach DR, Rabeneck L. 2009. Association of colonoscopy and death from colorectal cancer. Ann Intern Med 150(1):1-8.

Beagan K, McVey M. 2016. Linking DNA polymerase theta structure and function in health and disease. Cell Mol Life Sci 73(3):603-615.

Belkaid Y, Hand TW. 2014. Role of the microbiota in immunity and inflammation. Cell 157(1):121-141.

Benson AB, 3rd, Schrag D, Somerfield MR, Cohen AM, Figueredo AT, Flynn PJ, Krzyzanowska MK, Maroun J, McAllister P, Van Cutsem E, Brouwers M, Charette M, Haller DG. 2004. American Society of Clinical Oncology recommendations on adjuvant chemotherapy for stage II colon cancer. J Clin Oncol 22(16):3408-3419.

Bethke L, Webb E, Murray A, Schoemaker M, Johansen C, Christensen HC, Muir K, McKinney P, Hepworth S, Dimitropoulou P, Lophatananon A, Feychting M, Lonn S, Ahlbom A, Malmer B, Henriksson R, Auvinen A, Kiuru A, Salminen T, Swerdlow A, Houlston R. 2008. Comprehensive analysis of the role of DNA repair gene polymorphisms on risk of glioma. Hum Mol Genet 17(6):800-805.

Bhushan S, McLeod H, Walko CM. 2009. Role of pharmacogenetics as predictive biomarkers of response and/or toxicity in the treatment of colorectal cancer. Clin Colorectal Cancer 8(1):15-21.

Bosetti C, Levi F, Rosato V, Bertuccio P, Lucchini F, Negri E, La Vecchia C. 2011. Recent trends in colorectal cancer mortality in Europe. Int J Cancer 129(1):180-191.

Brandalize AP, Schuler-Faccini L, Hoffmann JS, Caleffi M, Cazaux C, Ashton-Prolla P. 2014. A DNA repair variant in POLQ (c.-1060A > G) is associated to hereditary breast cancer patients: a case-control study. BMC Cancer 14:850.

Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. 2018a. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin.

Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. 2018b. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 68(6):394-424.

Brenner H, Altenhofen L, Stock C, Hoffmeister M. 2013. Natural history of colorectal adenomas: birth cohort analysis among 3.6 million participants of screening colonoscopy. Cancer Epidemiol Biomarkers Prev 22(6):1043-1051.

Brenner H, Kloor M, Pox CP. 2014. Colorectal cancer. Lancet 383(9927):1490-1502.

Brondello JM, Pillaire MJ, Rodriguez C, Gourraud PA, Selves J, Cazaux C, Piette J. 2008. Novel evidences for a tumor suppressor role of Rev3, the catalytic subunit of Pol zeta. Oncogene 27(47):6093-6101.

Butterworth AS, Higgins JP, Pharoah P. 2006. Relative and absolute risk of colorectal cancer for individuals with a family history: a meta-analysis. Eur J Cancer 42(2):216-227.

Byrd JC, Bresalier RS. 2004. Mucins and mucin binding proteins in colorectal cancer. Cancer Metastasis Rev 23(1-2):77-99.

Calvert PM, Frucht H. 2002. The genetics of colorectal cancer. Ann Intern Med 137(7):603-612.

Cancer Genome Atlas N. 2012. Comprehensive molecular characterization of human colon and rectal cancer. Nature 487(7407):330-337.

Carethers JM, Jung BH. 2015. Genetics and Genetic Biomarkers in Sporadic Colorectal Cancer. Gastroenterology 149(5):1177-1190 e1173.

Carvalho FA, Aitken JD, Vijay-Kumar M, Gewirtz AT. 2012. Toll-like receptor-gut microbiota interactions: perturb at your own risk! Annu Rev Physiol 74:177-198.

Center MM, Jemal A, Smith RA, Ward E. 2009. Worldwide variations in colorectal cancer. CA Cancer J Clin 59(6):366-378.

Cipollini M, Figlioli G, Maccari G, Garritano S, De Santi C, Melaiu O, Barone E, Bambi F, Ermini S, Pellegrini G, Cristaudo A, Foddis R, Bonotti A, Romei C, Vivaldi A, Agate L, Molinari E, Barale R, Forsti A, Hemminki K, Elisei R, Gemignani F, Landi S. 2016. Polymorphisms within base and nucleotide excision repair pathways and risk of differentiated thyroid carcinoma. DNA Repair (Amst) 41:27-31.

Cipollini M, Landi S, Gemignani F. 2014. MicroRNA binding site polymorphisms as biomarkers in cancer management and research. Pharmgenomics Pers Med 7:173-191.

Claes AK, Zhou JY, Philpott DJ. 2015. NOD-Like Receptors: Guardians of Intestinal Mucosal Barriers. Physiology (Bethesda) 30(3):241-250.

Cooke GS, Campbell SJ, Bennett S, Lienhardt C, McAdam KP, Sirugo G, Sow O, Gustafson P, Mwangulu F, van Helden P, Fine P, Hoal EG, Hill AV. 2008. Mapping of a novel susceptibility locus suggests a role for MC3R and CTSZ in human tuberculosis. Am J Respir Crit Care Med 178(2):203-207.

Coppede F, Lopomo A, Spisni R, Migliore L. 2014. Genetic and epigenetic biomarkers for diagnosis, prognosis and treatment of colorectal cancer. World J Gastroenterol 20(4):943-956.

Crocco P, Montesanto A, Passarino G, Rose G. 2016. Polymorphisms Falling Within Putative miRNA Target Sites in the 3'UTR Region of SIRT2 and DRD2 Genes Are Correlated With Human Longevity. J Gerontol A Biol Sci Med Sci 71(5):586-592.

Cunningham D, Atkin W, Lenz HJ, Lynch HT, Minsky B, Nordlinger B, Starling N. 2010. Colorectal cancer. Lancet 375(9719):1030-1047.

Curtin NJ. 2012. DNA repair dysregulation from cancer driver to therapeutic target. Nat Rev Cancer 12(12):801-817.

Das V, Kalita J, Pal M. 2017. Predictive and prognostic biomarkers in colorectal cancer: A systematic review of recent advances and challenges. Biomed Pharmacother 87:8-19.

de la Chapelle A. 2004. Genetic predisposition to colorectal cancer. Nat Rev Cancer 4(10):769-780.

Dunlop MG, Dobbins SE, Farrington SM, Jones AM, Palles C, Whiffin N, Tenesa A, Spain S, Broderick P, Ooi LY, Domingo E, Smillie C, Henrion M, Frampton M, Martin L, Grimes G, Gorman M, Semple C, Ma YP, Barclay E, Prendergast J, Cazier JB, Olver B, Penegar S, Lubbe S, Chander I, Carvajal-Carmona LG, Ballereau S, Lloyd A, Vijayakrishnan J, Zgaga L, Rudan I, Theodoratou E, Colorectal Tumour Gene Identification C, Starr JM, Deary I, Kirac I, Kovacevic D, Aaltonen LA, Renkonen-Sinisalo L, Mecklin JP, Matsuda K, Nakamura Y, Okada Y, Gallinger S, Duggan DJ, Conti D, Newcomb P, Hopper J, Jenkins MA, Schumacher F, Casey G, Easton D, Shah M, Pharoah P, Lindblom A, Liu T, Swedish Low-Risk Colorectal Cancer Study G, Smith CG, West H, Cheadle JP, Group CC, Midgley R, Kerr DJ, Campbell H, Tomlinson IP, Houlston RS. 2012. Common variation near CDKN1A, POLD3 and SHROOM2 influences colorectal cancer risk. Nat Genet 44(7):770-776.

Dzikiewicz-Krawczyk A. 2008. The importance of making ends meet: mutations in genes and altered expression of proteins of the MRN complex and cancer. Mutat Res 659(3):262-273.

Edwards BK, Ward E, Kohler BA, Eheman C, Zauber AG, Anderson RN, Jemal A, Schymura MJ, Lansdorp-Vogelaar I, Seeff LC, van Ballegooijen M, Goede SL, Ries LA. 2010. Annual report to the nation on the status of cancer, 1975-2006, featuring colorectal cancer trends and impact of interventions (risk factors, screening, and treatment) to reduce future rates. Cancer 116(3):544-573.

Family L, Bensen JT, Troester MA, Wu MC, Anders CK, Olshan AF. 2015. Single-nucleotide polymorphisms in DNA bypass polymerase genes and association with breast cancer and breast cancer subtypes among African Americans and Whites. Breast Cancer Res Treat 149(1):181-190.

Fearon ER. 1995. Molecular genetics of colorectal cancer. Ann N Y Acad Sci 768:101-110.

Fearon ER, Vogelstein B. 1990. A genetic model for colorectal tumorigenesis. Cell 61(5):759-767.

Feng Y, Zhu J, Ou C, Deng Z, Chen M, Huang W, Li L. 2014. MicroRNA-145 inhibits tumour growth and metastasis in colorectal cancer by targeting fascin-1. Br J Cancer 110(9):2300-2309.

Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray F. 2015. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer 136(5):E359-386.

Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, Rosso S, Coebergh JW, Comber H, Forman D, Bray F. 2013. Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. Eur J Cancer 49(6):1374-1403.

Fernandez-Rozadilla C, Cazier JB, Moreno V, Crous-Bou M, Guino E, Duran G, Lamas MJ, Lopez R, Candamio S, Gallardo E, Pare L, Baiget M, Paez D, Lopez-Fernandez LA, Cortejoso L, Garcia MI, Bujanda L, Gonzalez D, Gonzalo V, Rodrigo L, Rene JM, Jover R, Brea-Fernandez A, Andreu M, Bessa X, Llor X, Xicola R, Palles C, Tomlinson I, Castellvi-Bel S, Castells A, Ruiz-Ponte C, Carracedo A, Consortium E. 2013a. Pharmacogenomics in colorectal cancer: a genome-wide association study to predict toxicity after 5-fluorouracil or FOLFOX administration. Pharmacogenomics J 13(3):209-217.

Fernandez-Rozadilla C, Cazier JB, Tomlinson IP, Carvajal-Carmona LG, Palles C, Lamas MJ, Baiget M, Lopez-Fernandez LA, Brea-Fernandez A, Abuli A, Bujanda L, Clofent J, Gonzalez D, Xicola R, Andreu M, Bessa X, Jover R, Llor X, Consortium E, Moreno V, Castells A, Carracedo A, Castellvi-Bel S, Ruiz-Ponte C. 2013b. A colorectal cancer genome-wide association study in a Spanish cohort identifies two variants associated with colorectal cancer risk at 1p33 and 8p12. BMC Genomics 14:55.

Fidler MM, Bray F, Vaccarella S, Soerjomataram I. 2017. Assessing global transitions in human development and colorectal cancer incidence. Int J Cancer 140(12):2709-2715.

Friedberg EC. 2003. DNA damage and repair. Nature 421(6921):436-440.

Fukata M, Vamadevan AS, Abreu MT. 2009. Toll-like receptors (TLRs) and Nod-like receptors (NLRs) in inflammatory disorders. Semin Immunol 21(4):242-253.

Gangadhar T, Schilsky RL. 2010. Molecular markers to individualize adjuvant therapy for colon cancer. Nat Rev Clin Oncol 7(6):318-325.

Gattolliat CH, Uguen A, Pesson M, Trillet K, Simon B, Doucet L, Robaszkiewicz M, Corcos L. 2015. MicroRNA and targeted mRNA expression profiling analysis in human colorectal adenomas and adenocarcinomas. Eur J Cancer 51(3):409-420.

Gonzalez-Pons M, Cruz-Correa M. 2015. Colorectal Cancer Biomarkers: Where Are We Now? Biomed Res Int 2015:149014.

Goodman JE, Mechanic LE, Luke BT, Ambs S, Chanock S, Harris CC. 2006. Exploring SNP-SNP interactions and colon cancer risk using polymorphism interaction analysis. Int J Cancer 118(7):1790-1797.

Grady WM, Markowitz SD. 2015. The molecular pathogenesis of colorectal cancer and its potential application to colorectal cancer screening. Dig Dis Sci 60(3):762-772.

Gupta BK, Maher DM, Ebeling MC, Sundram V, Koch MD, Lynch DW, Bohlmeyer T, Watanabe A, Aburatani H, Puumala SE, Jaggi M, Chauhan SC. 2012. Increased expression and aberrant localization of mucin 13 in metastatic colon cancer. J Histochem Cytochem 60(11):822-831.

Gustavsson B, Carlsson G, Machover D, Petrelli N, Roth A, Schmoll HJ, Tveit KM, Gibson F. 2015. A review of the evolution of systemic chemotherapy in the management of colorectal cancer. Clin Colorectal Cancer 14(1):1-10.

Hammond SM. 2015. An overview of microRNAs. Adv Drug Deliv Rev 87:3-14.

Hashim D, Boffetta P, La Vecchia C, Rota M, Bertuccio P, Malvezzi M, Negri E. 2016. The global decrease in cancer mortality: trends and disparities. Ann Oncol 27(5):926-933.

Hildrestrand GA, Neurauter CG, Diep DB, Castellanos CG, Krauss S, Bjoras M, Luna L. 2009. Expression patterns of Neil3 during embryonic brain development and neoplasia. BMC Neurosci 10:45.

Hoeijmakers JH. 2001. Genome maintenance mechanisms for preventing cancer. Nature 411(6835):366-374.

Hofer P, Hagmann M, Brezina S, Dolejsi E, Mach K, Leeb G, Baierl A, Buch S, Sutterluty-Fall H, Karner-Hanusch J, Bergmann MM, Bachleitner-Hofmann T, Stift A, Gerger A, Rotzer K, Karner J, Stattner S, Waldenberger M, Meitinger T, Strauch K, Linseisen J, Gieger C, Frommlet F, Gsur A. 2017. Bayesian and frequentist analysis of an Austrian genome-wide association study of colorectal cancer and advanced adenomas. Oncotarget 8(58):98623-98634.

Hoffe SE, Shridhar R, Biagioli MC. 2010. Radiation therapy for rectal cancer: current status and future directions. Cancer Control 17(1):25-34.

Hollingsworth MA, Swanson BJ. 2004. Mucins in cancer: protection and control of the cell surface. Nat Rev Cancer 4(1):45-60.

Hussain SK, Mu LN, Cai L, Chang SC, Park SL, Oh SS, Wang Y, Goldstein BY, Ding BG, Jiang Q, Rao J, You NC, Yu SZ, Papp JC, Zhao JK, Wang H, Zhang ZF. 2009. Genetic variation in immune regulation and DNA repair pathways and stomach cancer in China. Cancer Epidemiol Biomarkers Prev 18(8):2304-2309.

Huyghe JR, Bien SA, Harrison TA, Kang HM, Chen S, Schmit SL, Conti DV, Qu C, Jeon J, Edlund CK, Greenside P, Wainberg M, Schumacher FR, Smith JD, Levine DM, Nelson SC, Sinnott-Armstrong NA, Albanes D, Alonso MH, Anderson K, Arnau-Collell C, Arndt V, Bamia C, Banbury BL, Baron JA, Berndt SI, Bezieau S, Bishop DT, Boehm J, Boeing H, Brenner H, Brezina S, Buch S, Buchanan DD, Burnett-Hartman A, Butterbach K, Caan BJ, Campbell PT, Carlson CS, Castellvi-Bel S, Chan AT, Chang-Claude J, Chanock SJ, Chirlaque MD, Cho SH, Connolly CM, Cross AJ, Cuk K, Curtis KR, de la Chapelle A, Doheny KF, Duggan D, Easton DF, Elias SG, Elliott F, English DR, Feskens EJM, Figueiredo JC, Fischer R, FitzGerald LM, Forman D, Gala M, Gallinger S, Gauderman WJ, Giles GG, Gillanders E, Gong J, Goodman PJ, Grady WM, Grove JS, Gsur A, Gunter MJ, Haile RW, Hampe J, Hampel H, Harlid S, Hayes RB, Hofer P, Hoffmeister M, Hopper JL, Hsu WL, Huang WY, Hudson TJ, Hunter DJ, Ibanez-Sanz G, Idos GE, Ingersoll R, Jackson RD, Jacobs EJ, Jenkins MA, Joshi AD, Joshu CE, Keku TO, Key TJ, Kim HR, Kobayashi E, Kolonel LN, Kooperberg C, Kuhn T, Kury S, Kweon SS, Larsson SC, Laurie CA, Le Marchand L, Leal SM, Lee SC, Lejbkowicz F, Lemire M, Li CI, Li L, Lieb W, Lin Y, Lindblom A, Lindor NM, Ling H, Louie TL, Mannisto S, Markowitz SD, Martin V, Masala G, McNeil CE, Melas M, Milne RL, Moreno L, Murphy N, Myte R, Naccarati A, Newcomb PA, Offit K, Ogino S, Onland-Moret NC, Pardini B, Parfrey PS, Pearlman R, Perduca V, Pharoah PDP, Pinchev M, Platz EA, Prentice RL, Pugh E, Raskin L, Rennert G, Rennert HS, Riboli E, Rodriguez-Barranco M, Romm J, Sakoda LC, Schafmayer C, Schoen RE, Seminara D, Shah M, Shelford T, Shin MH, Shulman K, Sieri S, Slattery ML, Southey MC, Stadler ZK, Stegmaier C, Su YR, Tangen CM, Thibodeau SN, Thomas DC, Thomas SS, Toland AE, Trichopoulou A, Ulrich CM, Van Den Berg DJ, van Duijnhoven FJB, Van Guelpen B, van Kranen H, Vijai J, Visvanathan K, Vodicka P, Vodickova L, Vymetalkova V, Weigl K, Weinstein SJ, White E, Win AK, Wolf CR, Wolk A, Woods MO, Wu AH, Zaidi SH, Zanke BW, Zhang Q, Zheng W, Scacheri PC, Potter JD, Bassik MC, Kundaje A, Casey G, Moreno V, Abecasis GR, Nickerson DA, Gruber SB, Hsu L, Peters U. 2019. Discovery of common and rare genetic risk variants for colorectal cancer. Nat Genet 51(1):76-87.

Chen CH, Wang SW, Chen CW, Huang MR, Hung JS, Huang HC, Lin HH, Chen RJ, Shyu MK, Huang MC. 2013. MUC20 overexpression predicts poor prognosis and enhances EGF-

induced malignant phenotypes via activation of the EGFR-STAT3 pathway in endometrial cancer. Gynecol Oncol 128(3):560-567.

Cheng M, He B, Wan T, Zhu W, Han J, Zha B, Chen H, Yang F, Li Q, Wang W, Xu H, Ye T. 2012. 5-Fluorouracil nanoparticles inhibit hepatocellular carcinoma via activation of the p53 pathway in the orthotopic transplant mouse model. PLoS One 7(10):e47115.

Choudhury A, Elliott F, Iles MM, Churchman M, Bristow RG, Bishop DT, Kiltie AE. 2008. Analysis of variants in DNA damage signalling genes in bladder cancer. BMC Med Genet 9:69.

Chowdhury D, Choi YE, Brault ME. 2013. Charity begins at home: non-coding RNA functions in DNA repair. Nat Rev Mol Cell Biol 14(3):181-189.

Christmann M, Tomicic MT, Roos WP, Kaina B. 2003. Mechanisms of human DNA repair: an update. Toxicology 193(1-2):3-34.

Iorio MV, Croce CM. 2012. Causes and consequences of microRNA dysregulation. Cancer J 18(3):215-222.

Jasperson KW, Tuohy TM, Neklason DW, Burt RW. 2010. Hereditary and familial colon cancer. Gastroenterology 138(6):2044-2058.

Jemal A, Center MM, DeSantis C, Ward EM. 2010. Global patterns of cancer incidence and mortality rates and trends. Cancer Epidemiol Biomarkers Prev 19(8):1893-1907.

Johnson CM, Wei C, Ensor JE, Smolenski DJ, Amos CI, Levin B, Berry DA. 2013. Meta-analyses of colorectal cancer risk factors. Cancer Causes Control 24(6):1207-1222.

Karim-Kos HE, de Vries E, Soerjomataram I, Lemmens V, Siesling S, Coebergh JW. 2008. Recent trends of cancer in Europe: a combined approach of incidence, survival and mortality for 17 cancer sites since the 1990s. Eur J Cancer 44(10):1345-1389.

Karwacz K, Bricogne C, MacDonald D, Arce F, Bennett CL, Collins M, Escors D. 2011. PD-L1 co-stimulation contributes to ligand-induced T cell receptor down-modulation on CD8+ T cells. EMBO Mol Med 3(10):581-592.

Kataoka K, Shiraishi Y, Takeda Y, Sakata S, Matsumoto M, Nagano S, Maeda T, Nagata Y, Kitanaka A, Mizuno S, Tanaka H, Chiba K, Ito S, Watatani Y, Kakiuchi N, Suzuki H, Yoshizato T, Yoshida K, Sanada M, Itonaga H, Imaizumi Y, Totoki Y, Munakata W, Nakamura H, Hama N, Shide K, Kubuki Y, Hidaka T, Kameda T, Masuda K, Minato N, Kashiwase K, Izutsu K, Takaori-Kondo A, Miyazaki Y, Takahashi S, Shibata T, Kawamoto H, Akatsuka Y, Shimoda K, Takeuchi K, Seya T, Miyano S, Ogawa S. 2016. Aberrant PD-L1 expression through 3'-UTR disruption in multiple cancers. Nature 534(7607):402-406.

Kauffmann A, Rosselli F, Lazar V, Winnepenninckx V, Mansuet-Lupo A, Dessen P, van den Oord JJ, Spatz A, Sarasin A. 2008. High expression of DNA repair pathways is associated with metastasis in melanoma patients. Oncogene 27(5):565-573.

Kawamura K, Bahar R, Seimiya M, Chiyo M, Wada A, Okada S, Hatano M, Tokuhisa T, Kimura H, Watanabe S, Honda I, Sakiyama S, Tagawa M, J OW. 2004. DNA polymerase theta is preferentially expressed in lymphoid tissues and upregulated in human cancers. Int J Cancer 109(1):9-16.

Khurana E, Fu Y, Chakravarty D, Demichelis F, Rubin MA, Gerstein M. 2016. Role of non-coding sequence variants in cancer. Nat Rev Genet 17(2):93-108.

Kim YK, Shin JS, Nahm MH. 2016. NOD-Like Receptors in Infection, Immunity, and Diseases. Yonsei Med J 57(1):5-14.

Kitamoto S, Yamada N, Yokoyama S, Houjou I, Higashi M, Goto M, Batra SK, Yonezawa S. 2011. DNA methylation and histone H3-K9 modifications contribute to MUC17 expression. Glycobiology 21(2):247-256.

Knobel PA, Marti TM. 2011. Translesion DNA synthesis in the context of cancer research. Cancer Cell Int 11:39.

Kocarnik JM, Shiovitz S, Phipps AI. 2015. Molecular phenotypes of colorectal cancer and potential clinical applications. Gastroenterol Rep (Oxf) 3(4):269-276.

Kufe DW. 2009. Mucins in cancer: function, prognosis and therapy. Nat Rev Cancer 9(12):874-885.

Kuipers EJ, Grady WM, Lieberman D, Seufferlein T, Sung JJ, Boelens PG, van de Velde CJ, Watanabe T. 2015. Colorectal cancer. Nat Rev Dis Primers 1:15065.

La Vecchia C, Bosetti C, Lucchini F, Bertuccio P, Negri E, Boyle P, Levi F. 2010. Cancer mortality in Europe, 2000-2004, and an overview of trends since 1975. Ann Oncol 21(6):1323-1360.

Labelle M, Hynes RO. 2012. The initial hours of metastasis: the importance of cooperative host-tumor cell interactions during hematogenous dissemination. Cancer Discov 2(12):1091-1099.

Landi D, Gemignani F, Pardini B, Naccarati A, Garritano S, Vodicka P, Vodickova L, Canzian F, Novotny J, Barale R, Landi S. 2012. Identification of candidate genes carrying polymorphisms associated with the risk of colorectal cancer by analyzing the colorectal mutome and microRNAome. Cancer 118(19):4670-4680.

Lansdorp-Vogelaar I, van Ballegooijen M, Zauber AG, Habbema JD, Kuipers EJ. 2009. Effect of rising chemotherapy costs on the cost savings of colorectal cancer screening. J Natl Cancer Inst 101(20):1412-1422.

Lee JK, Chan AT. 2011. Molecular Prognostic and Predictive Markers in Colorectal Cancer: Current Status. Curr Colorectal Cancer Rep 7(2):136-144.

Lee PH, Shatkay H. 2008. F-SNP: computationally predicted functional SNPs for disease association studies. Nucleic Acids Res 36(Database issue):D820-824.

Lech G, Slotwinski R, Slodkowski M, Krasnodebski IW. 2016. Colorectal cancer tumour markers and biomarkers: Recent therapeutic advances. World J Gastroenterol 22(5):1745-1755.

Lemee F, Bergoglio V, Fernandez-Vidal A, Machado-Silva A, Pillaire MJ, Bieth A, Gentil C, Baker L, Martin AL, Leduc C, Lam E, Magdeleine E, Filleron T, Oumouhou N, Kaina B, Seki M, Grimal F, Lacroix-Triki M, Thompson A, Roche H, Bourdon JC, Wood RD, Hoffmann JS, Cazaux C. 2010. DNA polymerase theta up-regulation is associated with poor survival in breast cancer, perturbs DNA replication, and promotes genetic instability. Proc Natl Acad Sci U S A 107(30):13390-13395.

Li KL, A.; Kandarian, B.S.; Manugian, V.; Peterson, M.R.; Batra, S.K.; Ho, S.B. 2011. Aberrant Expression of MUC17 Correlates With Survival in Advanced Stage Colorectal Adenocarcinomas. Gastroenterology 140(5):S-345.

Li WQ, Hu N, Hyland PL, Gao Y, Wang ZM, Yu K, Su H, Wang CY, Wang LM, Chanock SJ, Burdett L, Ding T, Qiao YL, Fan JH, Wang Y, Xu Y, Shi JX, Gu F, Wheeler W, Xiong XQ, Giffen C, Tucker MA, Dawsey SM, Freedman ND, Abnet CC, Goldstein AM, Taylor PR. 2013. Genetic variants in DNA repair pathway genes and risk of esophageal squamous cell carcinoma and gastric adenocarcinoma in a Chinese population. Carcinogenesis 34(7):1536-1542.

Lieberman DA, Rex DK, Winawer SJ, Giardiello FM, Johnson DA, Levin TR. 2012. Guidelines for colonoscopy surveillance after screening and polypectomy: a consensus update by the US Multi-Society Task Force on Colorectal Cancer. Gastroenterology 143(3):844-857.

Liu R, Truax AD, Chen L, Hu P, Li Z, Chen J, Song C, Chen L, Ting JP. 2015. Expression profile of innate immune receptors, NLRs and AIM2, in human colorectal cancer: correlation with cancer stages and inflammasome components. Oncotarget 6(32):33456-33469.

Longley DB, Harkin DP, Johnston PG. 2003. 5-fluorouracil: mechanisms of action and clinical strategies. Nat Rev Cancer 3(5):330-338.

Lu S, Catalano C, Huhn S, Pardini B, Partu L, Vymetalkova V, Vodickova L, Levy M, Buchler T, Hemminki K, Vodicka P, Forsti A. 2019. Single nucleotide polymorphisms within MUC4 are associated with colorectal cancer survival. PLoS One 14(5):e0216666.

Lu S, Pardini B, Cheng B, Naccarati A, Huhn S, Vymetalkova V, Vodickova L, Buchler T, Hemminki K, Vodicka P, Forsti A. 2014. Single nucleotide polymorphisms within interferon signaling pathway genes are associated with colorectal cancer susceptibility and survival. PLoS One 9(10):e111061.

Lu Y, Yang H, Yuan L, Liu G, Zhang C, Hong M, Liu Y, Zhou M, Chen F, Li X. 2016. Overexpression of miR-335 confers cell proliferation and tumour growth to colorectal carcinoma cells. Mol Cell Biochem 412(1-2):235-245.

Lupfer C, Kanneganti TD. 2013. Unsolved Mysteries in NLR Biology. Front Immunol 4:285.

Lynch D, Murphy A. 2016. The emerging role of immunotherapy in colorectal cancer. Ann Transl Med 4(16):305.

Lynch HT, de la Chapelle A. 2003. Hereditary colorectal cancer. N Engl J Med 348(10):919-932.

Macha MA, Krishn SR, Jahan R, Banerjee K, Batra SK, Jain M. 2015. Emerging potential of natural products for targeting mucins for therapy against inflammation and cancer. Cancer Treat Rev 41(3):277-288.

Marin F, Bonet C, Munoz X, Garcia N, Pardo ML, Ruiz-Liso JM, Alonso P, Capella G, Sanz-Anquela JM, Gonzalez CA, Sala N. 2012. Genetic variation in MUC1, MUC2 and MUC6 genes and evolution of gastric cancer precursor lesions in a long-term follow-up in a high-risk area in Spain. Carcinogenesis 33(5):1072-1080.

Megnin-Chanet F, Bollet MA, Hall J. 2010. Targeting poly(ADP-ribose) polymerase activity for cancer therapy. Cell Mol Life Sci 67(21):3649-3662.

Meissner TB, Li A, Biswas A, Lee KH, Liu YJ, Bayir E, Iliopoulos D, van den Elsen PJ, Kobayashi KS. 2010. NLR family member NLRC5 is a transcriptional regulator of MHC class I genes. Proc Natl Acad Sci U S A 107(31):13794-13799.

Mosor M, Ziolkowska I, Pernak-Schwarz M, Januszkiewicz-Lewandowska D, Nowak J. 2006. Association of the heterozygous germline I171V mutation of the NBS1 gene with childhood acute lymphoblastic leukemia. Leukemia 20(8):1454-1456.

Murphy N, Moreno V, Hughes DJ, Vodicka L, Vodicka P, Aglago EK, Gunter MJ, Jenab M. 2019. Lifestyle and dietary environmental factors in colorectal cancer susceptibility. Mol Aspects Med.

Naccarati A, Pardini B, Hemminki K, Vodicka P. 2007. Sporadic colorectal cancer and individual susceptibility: a review of the association studies investigating the role of DNA repair genetic polymorphisms. Mutat Res 635(2-3):118-145.

Naccarati A, Pardini B, Stefano L, Landi D, Slyskova J, Novotny J, Levy M, Polakova V, Lipska L, Vodicka P. 2012. Polymorphisms in miRNA-binding sites of nucleotide excision repair genes and colorectal cancer risk. Carcinogenesis 33(7):1346-1351.

Neerincx A, Castro W, Guarda G, Kufer TA. 2013. NLRC5, at the Heart of Antigen Presentation. Front Immunol 4:397.

Nikolouzakis TK, Vassilopoulou L, Fragkiadaki P, Mariolis Sapsakos T, Papadakis GZ, Spandidos DA, Tsatsakis AM, Tsiaoussis J. 2018. Improving diagnosis, prognosis and prediction by using biomarkers in CRC patients (Review). Oncol Rep 39(6):2455-2472.

Oviedo-Boyso J, Bravo-Patino A, Baizabal-Aguirre VM. 2014. Collaborative action of Toll-like and NOD-like receptors as modulators of the inflammatory response to pathogenic bacteria. Mediators Inflamm 2014:432785.

Pan J, Chi P, Lu X, Xu Z. 2012. Genetic polymorphisms in translesion synthesis genes are associated with colorectal cancer risk and metastasis in Han Chinese. Gene 504(2):151-155.

Pandiyan P, Hegel JK, Krueger M, Quandt D, Brunner-Weinzierl MC. 2007. High IFN-gamma production of individual CD8 T lymphocytes is controlled by CD152 (CTLA-4). J Immunol 178(4):2132-2140.

Pardini B, Corrado A, Paolicchi E, Cugliari G, Berndt SI, Bezieau S, Bien SA, Brenner H, Caan BJ, Campbell PT, Casey G, Chan AT, Chang-Claude J, Cotterchio M, Gala M, Gallinger SJ, Haile RW, Harrison TA, Hayes RB, Hoffmeister M, Hopper JL, Hsu L, Huyghe J, Jenkins MA, Le Marchand L, Lin Y, Lindor NM, Nan H, Newcomb PA, Ogino S, Potter JD, Schoen RE, Slattery ML, White E, Vodickova L, Vymetalkova V, Vodicka P, Gemignani F, Peters U, Naccarati A, Landi S. 2019. DNA repair and cancer in colon and rectum: Novel players in genetic susceptibility. Int J Cancer.

Pardini B, Naccarati A, Novotny J, Smerhovsky Z, Vodickova L, Polakova V, Hanova M, Slyskova J, Tulupova E, Kumar R, Bortlik M, Barale R, Hemminki K, Vodicka P. 2008. DNA repair genetic polymorphisms and risk of colorectal cancer in the Czech Republic. Mutat Res 638(1-2):146-153.

Pardini B, Rosa F, Barone E, Di Gaetano C, Slyskova J, Novotny J, Levy M, Garritano S, Vodickova L, Buchler T, Gemignani F, Landi S, Vodicka P, Naccarati A. 2013. Variation within 3'-UTRs of base excision repair genes and response to therapy in colorectal cancer patients: A potential modulation of microRNAs binding. Clin Cancer Res 19(21):6044-6056.

Passardi A, Canale M, Valgiusti M, Ulivi P. 2017. Immune Checkpoints as a Target for Colorectal Cancer Treatment. Int J Mol Sci 18(6).

Pavelitz T, Renfro L, Foster NR, Caracol A, Welsch P, Lao VV, Grady WB, Niedzwiecki D, Saltz LB, Bertagnolli MM, Goldberg RM, Rabinovitch PS, Emond M, Monnat RJ, Jr., Maizels N. 2014. MRE11-deficiency associated with improved long-term disease free survival and overall survival in a subset of stage III colon cancer patients in randomized CALGB 89803 trial. PLoS One 9(10):e108483.

Peltomaki P. 2001. Deficient DNA mismatch repair: a common etiologic factor for colon cancer. Hum Mol Genet 10(7):735-740.

Peters U, Bien S, Zubair N. 2015. Genetic architecture of colorectal cancer. Gut 64(10):1623-1636.

Peters U, Jiao S, Schumacher FR, Hutter CM, Aragaki AK, Baron JA, Berndt SI, Bezieau S, Brenner H, Butterbach K, Caan BJ, Campbell PT, Carlson CS, Casey G, Chan AT, Chang-Claude J, Chanock SJ, Chen LS, Coetzee GA, Coetzee SG, Conti DV, Curtis KR, Duggan D,

Edwards T, Fuchs CS, Gallinger S, Giovannucci EL, Gogarten SM, Gruber SB, Haile RW, Harrison TA, Hayes RB, Henderson BE, Hoffmeister M, Hopper JL, Hudson TJ, Hunter DJ, Jackson RD, Jee SH, Jenkins MA, Jia WH, Kolonel LN, Kooperberg C, Kury S, Lacroix AZ, Laurie CC, Laurie CA, Le Marchand L, Lemire M, Levine D, Lindor NM, Liu Y, Ma J, Makar KW, Matsuo K, Newcomb PA, Potter JD, Prentice RL, Qu C, Rohan T, Rosse SA, Schoen RE, Seminara D, Shrubsole M, Shu XO, Slattery ML, Taverna D, Thibodeau SN, Ulrich CM, White E, Xiang Y, Zanke BW, Zeng YX, Zhang B, Zheng W, Hsu L, Colon Cancer Family R, the G, Epidemiology of Colorectal Cancer C. 2013. Identification of Genetic Susceptibility Loci for Colorectal Tumors in a Genome-Wide Meta-analysis. Gastroenterology 144(4):799-807 e724.

Pillaire MJ, Selves J, Gordien K, Gourraud PA, Gentil C, Danjoux M, Do C, Negre V, Bieth A, Guimbaud R, Trouche D, Pasero P, Mechali M, Hoffmann JS, Cazaux C. 2010. A 'DNA replication' signature of progression and negative outcome in colorectal cancer. Oncogene 29(6):876-887.

Preskill C, Weidhaas JB. 2013. SNPs in microRNA binding sites as prognostic and predictive cancer biomarkers. Crit Rev Oncog 18(4):327-340.

Quasar Collaborative G, Gray R, Barnwell J, McConkey C, Hills RK, Williams NS, Kerr DJ. 2007. Adjuvant chemotherapy versus observation in patients with colorectal cancer: a randomised study. Lancet 370(9604):2020-2029.

Ramzy I, Hasaballah M, Marzaban R, Shaker O, Soliman ZA. 2015. Evaluation of microRNAs-29a, 92a and 145 in colorectal carcinoma as candidate diagnostic markers: An Egyptian pilot study. Clin Res Hepatol Gastroenterol 39(4):508-515.

Raskov H, Pommergaard HC, Burcharth J, Rosenberg J. 2014. Colorectal carcinogenesis--update and perspectives. World J Gastroenterol 20(48):18151-18164.

Rendleman J, Antipin Y, Reva B, Adaniel C, Przybylo JA, Dutra-Clarke A, Hansen N, Heguy A, Huberman K, Borsu L, Paltiel O, Ben-Yehuda D, Brown JR, Freedman AS, Sander C, Zelenetz A, Klein RJ, Shao Y, Lacher M, Vijai J, Offit K, Kirchhoff T. 2014. Genetic variation in DNA repair pathways and risk of non-Hodgkin's lymphoma. PLoS One 9(7):e101685.

Riella LV, Paterson AM, Sharpe AH, Chandraker A. 2012. Role of the PD-1 pathway in the immune response. Am J Transplant 12(10):2575-2587.

Roos WP, Thomas AD, Kaina B. 2016. DNA damage and the balance between survival and death in cancer biology. Nat Rev Cancer 16(1):20-33.

Ryan BM, Robles AI, McClary AC, Haznadar M, Bowman ED, Pine SR, Brown D, Khan M, Shiraishi K, Kohno T, Okayama H, Modali R, Yokota J, Harris CC. 2015. Identification of a functional SNP in the 3'UTR of CXCR2 that is associated with reduced risk of lung cancer. Cancer Res 75(3):566-575.

Samadder NJ, Jasperson K, Burt RW. 2015. Hereditary and common familial colorectal cancer: evidence for colorectal screening. Dig Dis Sci 60(3):734-747.

Sargent DJ, Patiyil S, Yothers G, Haller DG, Gray R, Benedetti J, Buyse M, Labianca R, Seitz JF, O'Callaghan CJ, Francini G, Grothey A, O'Connell M, Catalano PJ, Kerr D, Green E, Wieand HS, Goldberg RM, de Gramont A, Group A. 2007. End points for colon cancer adjuvant trials: observations and recommendations based on individual patient data from 20,898 patients enrolled onto 18 randomized trials from the ACCENT Group. J Clin Oncol 25(29):4569-4574.

Sethupathy P, Collins FS. 2008. MicroRNA target site polymorphisms and human disease. Trends Genet 24(10):489-497.

Shafee TL, R. 2017. Eukaryotic and prokaryotic gene structure. WikiJournal of Medicine 4(1).

Sharma V, Misteli T. 2013. Non-coding RNAs in DNA damage and repair. FEBS Lett 587(13):1832-1839.

Sharp RC, Abdulrahim M, Naser ES, Naser SA. 2015. Genetic Variations of PTPN2 and PTPN22: Role in the Pathogenesis of Type 1 Diabetes and Crohn's Disease. Front Cell Infect Microbiol 5:95.

Shinmura K, Kato H, Kawanishi Y, Igarashi H, Goto M, Tao H, Inoue Y, Nakamura S, Misawa K, Mineta H, Sugimura H. 2016. Abnormal Expressions of DNA Glycosylase Genes NEIL1, NEIL2, and NEIL3 Are Associated with Somatic Mutation Loads in Human Cancer. Oxid Med Cell Longev 2016:1546392.

Schreuders EH, Ruco A, Rabeneck L, Schoen RE, Sung JJ, Young GP, Kuipers EJ. 2015. Colorectal cancer screening: a global overview of existing programmes. Gut 64(10):1637-1649.

Sidransky D. 2002. Emerging molecular markers of cancer. Nat Rev Cancer 2(3):210-219.

Siegel RL, Miller KD, Fedewa SA, Ahnen DJ, Meester RGS, Barzi A, Jemal A. 2017. Colorectal cancer statistics, 2017. CA Cancer J Clin 67(3):177-193.

Simon K. 2016. Colorectal cancer development and advances in screening. Clin Interv Aging 11:967-976.

Simpson JA, Al-Attar A, Watson NF, Scholefield JH, Ilyas M, Durrant LG. 2010. Intratumoral T cell infiltration, MHC class I and STAT1 as biomarkers of good prognosis in colorectal cancer. Gut 59(7):926-933.

Slaby O. 2016. Non-coding RNAS in colorectal cancer. New York, NY: Springer Berlin Heidelberg. pages cm p.

Slaby O, Bienertova-Vasku J, Svoboda M, Vyzula R. 2012. Genetic polymorphisms and microRNAs: new direction in molecular epidemiology of solid cancer. J Cell Mol Med 16(1):8-21.

Sladek R, Rocheleau G, Rung J, Dina C, Shen L, Serre D, Boutin P, Vincent D, Belisle A, Hadjadj S, Balkau B, Heude B, Charpentier G, Hudson TJ, Montpetit A, Pshezhetsky AV, Prentki M, Posner BI, Balding DJ, Meyre D, Polychronakos C, Froguel P. 2007. A genome-wide association study identifies novel risk loci for type 2 diabetes. Nature 445(7130):881-885.

Slyskova J, Lorenzo Y, Karlsen A, Carlsen MH, Novosadova V, Blomhoff R, Vodicka P, Collins AR. 2014. Both genetic and dietary factors underlie individual differences in DNA damage levels and DNA repair capacity. DNA Repair (Amst) 16:66-73.

Stallons LJ, McGregor WG. 2010. Translesion synthesis polymerases in the prevention and promotion of carcinogenesis. J Nucleic Acids 2010.

Stracci F, Zorzi M, Grazzini G. 2014. Colorectal cancer screening: tests, strategies, and perspectives. Front Public Health 2:210.

Suchanek S, Grega T, Zavoral M. Colorectal cancer screening. 2018. Vnitr Lek 64(6):679-683.

Tan SC. 2018. Low penetrance genetic polymorphisms as potential biomarkers for colorectal cancer predisposition. J Gene Med 20(4):e3010.

Teo MT, Dyrskjot L, Nsengimana J, Buchwald C, Snowden H, Morgan J, Jensen JB, Knowles MA, Taylor G, Barrett JH, Borre M, Orntoft TF, Bishop DT, Kiltie AE. 2014. Next-generation sequencing identifies germline MRE11A variants as markers of radiotherapy outcomes in muscle-invasive bladder cancer. Ann Oncol 25(4):877-883.

Teo MT, Landi D, Taylor CF, Elliott F, Vaslin L, Cox DG, Hall J, Landi S, Bishop DT, Kiltie AE. 2012. The role of microRNA-binding site polymorphisms in DNA repair genes as risk factors for bladder cancer and breast cancer and their impact on radiotherapy outcomes. Carcinogenesis 33(3):581-586.

Therkildsen C, Bergmann TK, Henrichsen-Schnack T, Ladelund S, Nilbert M. 2014. The predictive value of KRAS, NRAS, BRAF, PIK3CA and PTEN for anti-EGFR treatment in metastatic colorectal cancer: A systematic review and meta-analysis. Acta Oncol 53(7):852-864.

Tomlinson IP, Carvajal-Carmona LG, Dobbins SE, Tenesa A, Jones AM, Howarth K, Palles C, Broderick P, Jaeger EE, Farrington S, Lewis A, Prendergast JG, Pittman AM, Theodoratou E, Olver B, Walker M, Penegar S, Barclay E, Whiffin N, Martin L, Ballereau S, Lloyd A, Gorman M, Lubbe S, Consortium C, Collaborators C, Consortium E, Howie B, Marchini J, Ruiz-Ponte C, Fernandez-Rozadilla C, Castells A, Carracedo A, Castellvi-Bel S, Duggan D, Conti D, Cazier JB, Campbell H, Sieber O, Lipton L, Gibbs P, Martin NG, Montgomery GW, Young J, Baird PN, Gallinger S, Newcomb P, Hopper J, Jenkins MA, Aaltonen LA, Kerr DJ, Cheadle J, Pharoah P, Casey G, Houlston RS, Dunlop MG. 2011. Multiple common susceptibility variants near BMP pathway loci GREM1, BMP4, and BMP2 explain part of the missing heritability of colorectal cancer. PLoS Genet 7(6):e1002105.

Tougeron D, Mouillet G, Trouilloud I, Lecomte T, Coriat R, Aparicio T, Des Guetz G, Lecaille C, Artru P, Sickersen G, Cauchin E, Sefrioui D, Boussaha T, Ferru A, Matysiak-Budnik T, Silvain C, Karayan-Tapon L, Pages JC, Vernerey D, Bonnetain F, Michel P, Taieb J, Zaanan A. 2016. Efficacy of Adjuvant Chemotherapy in Colon Cancer With Microsatellite Instability: A Large Multicenter AGEO Study. J Natl Cancer Inst 108(7).

Twelves C, Wong A, Nowacki MP, Abt M, Burris H, 3rd, Carrato A, Cassidy J, Cervantes A, Fagerberg J, Georgoulias V, Husseini F, Jodrell D, Koralewski P, Kroning H, Maroun J, Marschner N, McKendrick J, Pawlicki M, Rosso R, Schuller J, Seitz JF, Stabuc B, Tujakowski J, Van Hazel G, Zaluski J, Scheithauer W. 2005. Capecitabine as adjuvant treatment for stage III colon cancer. N Engl J Med 352(26):2696-2704.

Vacante M, Borzi AM, Basile F, Biondi A. 2018. Biomarkers in colorectal cancer: Current clinical utility and future perspectives. World J Clin Cases 6(15):869-881.

Valle L, de Voer RM, Goldberg Y, Sjursen W, Forsti A, Ruiz-Ponte C, Caldes T, Garre P, Olsen MF, Nordling M, Castellvi-Bel S, Hemminki K. 2019. Update on genetic predisposition to colorectal cancer and polyposis. Mol Aspects Med.

Van Der Kraak L, Goel G, Ramanan K, Kaltenmeier C, Zhang L, Normolle DP, Freeman GJ, Tang D, Nason KS, Davison JM, Luketich JD, Dhupar R, Lotze MT. 2016. 5-Fluorouracil upregulates cell surface B7-H1 (PD-L1) expression in gastrointestinal cancers. J Immunother Cancer 4:65.

van Putten JPM, Strijbis K. 2017. Transmembrane Mucins: Signaling Receptors at the Intersection of Inflammation and Cancer. J Innate Immun 9(3):281-299.

Varadi V, Bevier M, Grzybowska E, Johansson R, Enquist K, Henriksson R, Butkiewicz D, Pamula-Pilat J, Tecza K, Hemminki K, Lenner P, Forsti A. 2011. Genetic variation in genes encoding for polymerase zeta subunits associates with breast cancer risk, tumour characteristics and survival. Breast Cancer Res Treat 129(1):235-245.

Verschuren JJ, Trompet S, Deelen J, Stott DJ, Sattar N, Buckley BM, Ford I, Heijmans BT, Guchelaar HJ, Houwing-Duistermaat JJ, Slagboom PE, Jukema JW. 2013. Non-homologous end-joining pathway associated with occurrence of myocardial infarction: gene set analysis of genome-wide association study data. PLoS One 8(2):e56262.

Vineis P, Manuguerra M, Kavvoura FK, Guarrera S, Allione A, Rosa F, Di Gregorio A, Polidoro S, Saletta F, Ioannidis JP, Matullo G. 2009. A field synopsis on low-penetrance variants in DNA repair genes and cancer susceptibility. J Natl Cancer Inst 101(1):24-36.

Vishnubalaji R, Hamam R, Abdulla MH, Mohammed MA, Kassem M, Al-Obeed O, Aldahmash A, Alajez NM. 2015. Genome-wide mRNA and miRNA expression profiling reveal multiple regulatory networks in colorectal cancer. Cell Death Dis 6:e1614.

Vlad AM, Diaconu I, Gantt KR. 2006. MUC1 in endometriosis and ovarian cancer. Immunol Res 36(1-3):229-236.

Vodicka P, Stetina R, Polakova V, Tulupova E, Naccarati A, Vodickova L, Kumar R, Hanova M, Pardini B, Slyskova J, Musak L, De Palma G, Soucek P, Hemminki K. 2007. Association of DNA repair polymorphisms with DNA repair functional outcomes in healthy human subjects. Carcinogenesis 28(3):657-664.

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

Wang YP, Song GH, Chen J, Xiao C, Li C, Zhong L, Sun X, Wang ZW, Deng GL, Yu FD, Xue YM, Tang HM, Peng ZH, Wang XL. 2016. Elevated OCT1 participates in colon tumorigenesis and independently predicts poor prognoses of colorectal cancer patients. Tumour Biol 37(3):3247-3255.

Watson NF, Ramage JM, Madjd Z, Spendlove I, Ellis IO, Scholefield JH, Durrant LG. 2006. Immunosurveillance is active in colorectal cancer as downregulation but not complete loss of MHC class I expression correlates with a poor prognosis. Int J Cancer 118(1):6-10.

Williams GJ, Lees-Miller SP, Tainer JA. 2010. Mre11-Rad50-Nbs1 conformations and the control of sensing, signaling, and effector responses at DNA double-strand breaks. DNA Repair (Amst) 9(12):1299-1306.

Wilmanski JM, Petnicki-Ocwieja T, Kobayashi KS. 2008. NLR proteins: integral members of innate immunity and mediators of inflammatory diseases. J Leukoc Biol 83(1):13-30.

Witvliet DK, Strokach A, Giraldo-Forero AF, Teyra J, Colak R, Kim PM. 2016. ELASPIC web-server: proteome-wide structure-based prediction of mutation effects on protein stability and binding affinity. Bioinformatics 32(10):1589-1591.

Woenckhaus M, Merk J, Stoehr R, Schaeper F, Gaumann A, Wiebe K, Hartmann A, Hofstaedter F, Dietmaier W. 2008. Prognostic value of FHIT, CTNNB1, and MUC1 expression in non-small cell lung cancer. Hum Pathol 39(1):126-136.

Wolf AMD, Fontham ETH, Church TR, Flowers CR, Guerra CE, LaMonte SJ, Etzioni R, McKenna MT, Oeffinger KC, Shih YT, Walter LC, Andrews KS, Brawley OW, Brooks D, Fedewa SA, Manassaram-Baptiste D, Siegel RL, Wender RC, Smith RA. 2018. Colorectal cancer screening for average-risk adults: 2018 guideline update from the American Cancer Society. CA Cancer J Clin 68(4):250-281.

Wu Z, Wang P, Song C, Wang K, Yan R, Li J, Dai L. 2015. Evaluation of miRNA-binding-site SNPs of MRE11A, NBS1, RAD51 and RAD52 involved in HRR pathway genes and risk of breast cancer in China. Mol Genet Genomics 290(3):1141-1153.

Xi T, Jones IM, Mohrenweiser HW. 2004. Many amino acid substitution variants identified in DNA repair genes during human population screenings are predicted to impact protein function. Genomics 83(6):970-979.

Xiao X, Wang L, Wei P, Chi Y, Li D, Wang Q, Ni S, Tan C, Sheng W, Sun M, Zhou X, Du X. 2013. Role of MUC20 overexpression as a predictor of recurrence and poor outcome in colorectal cancer. J Transl Med 11:151.

Xu Y, Pasche B. 2007. TGF-beta signaling alterations and susceptibility to colorectal cancer. Hum Mol Genet 16 Spec No 1:R14-20.

Yi Y, Kamata-Sakurai M, Denda-Nagai K, Itoh T, Okada K, Ishii-Schrade K, Iguchi A, Sugiura D, Irimura T. 2010. Mucin 21/epiglycanin modulates cell adhesion. J Biol Chem 285(28):21233-21240.

Yonezawa S, Goto M, Yamada N, Higashi M, Nomoto M. 2008a. Expression profiles of MUC1, MUC2, and MUC4 mucins in human neoplasms and their relationship with biological behavior. Proteomics 8(16):3329-3341.

Yonezawa S, Higashi M, Yamada N, Goto M. 2008b. Precursor lesions of pancreatic cancer. Gut Liver 2(3):137-154.

Yonezawa S, Higashi M, Yamada N, Yokoyama S, Kitamoto S, Kitajima S, Goto M. 2011. Mucins in human neoplasms: clinical pathology, gene expression and diagnostic application. Pathol Int 61(12):697-716.

Yoshihama S, Roszik J, Downs I, Meissner TB, Vijayan S, Chapuy B, Sidiq T, Shipp MA, Lizee GA, Kobayashi KS. 2016. NLRC5/MHC class I transactivator is a target for immune evasion in cancer. Proc Natl Acad Sci U S A 113(21):5999-6004.

Zhan T, Rindtorff N, Boutros M. 2017. Wnt signaling in cancer. Oncogene 36(11):1461-1473.

Zheng F, Yu H, Lu J. 2019. High expression of MUC20 drives tumorigenesis and predicts poor survival in endometrial cancer. J Cell Biochem.

Zhong Y, Kinio A, Saleh M. 2013. Functions of NOD-Like Receptors in Human Diseases. Front Immunol 4:333.

Ziegler A, Koch A, Krockenberger K, Grosshennig A. 2012. Personalized medicine using DNA biomarkers: a review. Hum Genet 131(10):1627-1638.

Ziolkowska-Suchanek I, Mosor M, Wierzbicka M, Rydzanicz M, Baranowska M, Nowak J. 2013. The MRN protein complex genes: MRE11 and RAD50 and susceptibility to head and neck cancers. Mol Cancer 12(1):113.

8. MANUSCRIPTS I-VII *in extenso*