Charles University

Faculty of Medicine in Pilsen

Anatomy, Histology and Embryology

Molecular prognostic and predictive markers in colorectal cancer

Dissertation

Supervisor:

Dr. Pavel Pitule

Consultant:

Res. Prof. Pavel Vodicka

Pilsen *2023*   **Nazila Navvabi**

ABSTRACT

Colorectal cancer is currently one of the three major causes of cancer-related death worldwide. Despite the fact that the incidence of CRC has been regularly rising globally over the last few decades, mainly in developing countries, the overall survival rate has increased given the advances in targeted agents and cytotoxic chemotherapy. However, patient survival is highly dependent on the tumor stage at the time of diagnosis.

Thus, in order to help clinicians to treat colorectal cancer, it is necessary to introduce more effective tools that will improve not only early diagnosis, but also prediction of the progression of the disease and response to chemotherapy. In that way, they will be able to decrease both morbidity and mortality of their patients. In accordance with that, colon cancer research has described numerous biomarkers for diagnostic, prognostic and predictive purposes that either alone or as part of a panel would help improve patient's clinical management. The thesis introduction describes the proposed CRC biomarkers divided based on the clinical specimen that is examined. The new insight in CRC monitoring is discussed, presenting promising emerging biomarkers (telomerase activity, telomere length and micronuclei frequency). In relation to my research, the thesis also focuses on alternative splicing that has emerged as an important regulator and potential treatment target in CRC. Alterations in gene expression leading to colorectal carcinogenesis results in dysregulated levels of nucleic acids and proteins, which can be utilized for the identification of modern, minimally invasive molecular biomarkers. Alternative splicing factors regulate alternative exon repression or activation which allows one pre-mRNA to code for multiple protein isoforms. One of the alternative splicing factors are MBNL proteins, a family of three members - MBNL1, MBNL2 and MBNL3 - that was primarily linked with the development of myotonic dystrophy.

The goal of my work that is presented here was to analyze gene expression of the MBNL family of regulators of alternative splicing in various stages of colorectal cancer development, together with the MBNL-target splicing events in FOXP1 and EPB41L3 genes and tumor-related CD44 variants.

The study was done using samples of tumor tissue and non-malignant mucosa from 108 CRC patients. After RNA isolation and reverse transcription, the relative gene

expression of a selected gene panel was tested by quantitative real-time PCR, followed by statistical analysis.

We observed that MBNL expression was decreased in tumor tissue compared to non-tumor mucosa. In addition, lower expression was observed for the variants of FOXP1 and EPB41L3, while higher expression in tumor tissue was detected both for total CD44 and its cancer-related variants 3 and 6. Transcript levels of the MBNL genes were not found to be related to any of the studied clinicopathological characteristics. Multiple significant associations were identified in the target gene panel, including higher transcript levels of FOXP1 and CD44v3 in patients with distant metastases and connections between recurrence-free survival and altered levels of FOXP1 and CD44v3.

Our results identified for the first-time deregulation of MBNL genes in colorectal cancer. Given the physiological role of MBNL proteins, we hypothesized that down-regulation of their transcripts in tumor tissue compared to matched non-tumor mucosa can lead to transition of alternative splicing patterns towards a less differentiated phenotype, that consequentially induce further tumor development. We believe that our research highlights the importance of alternative splicing regulation for tumor growth and propagation. Furthermore, an utilisation of biomarkers involved in the repair of oxidative DNA damage and in epigenetics in CRC patients is discussed as well.

Preface and statement

I dedicate my dissertation primarily to my greatest Dad, Mr. Ali Navvabi, who donated his life to me and my sister. He was always looking forward to this moment to see my achievement in study and life. I am sure he is alive and he is seeing my achievement… Dad, you are always near me, in my heart forever. I love you my great Dad….

Declaration

I declare hereby that I completed this dissertation thesis by myself and that I mentioned and cited properly all the sources and literature. At the same time, I declare that this thesis was not used to obtain another or the same title.

I agree with permanent deposition of an electronic version of my thesis in the system database of interuniversity project Thesis.cz for a permanent control of similarities of theses.

In Pilsen,

<div align="right">Nazila Navvabi</div>

Table of Contents

1 Abbreviations

AA	Alternative acceptor
ACTB	Actin beta
APC	Adenomatous polyposis coli
AS	Alternative splicing
ASCO	American Society of Clinical Oncology
ASS	Alternative splicing sites
BLAST	Basic Local Alignment Search Tool
CDH17	Cadherin 17
cDNA	Complementary DNA
CDX2	Caudal type homeobox 2
CEA	Carcinoembryonic antigen
CF	Cell-free
cfDNA	Circulating cell-free DNA
ctDNA	Circulating tumour DNA
CIN	Chromosomal instability
CIMP	CpG island methylator phenotype
CKs	Cytokeratins
CpG	Cytosine preceding guanine
CRC	Colorectal cancer
DPEP1	Dipeptidase 1
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
DSB	Double strand-breaks
EGFR	Epidermal growth factor receptor
EMT	Epithelial-to-mesenchymal transition
EGF	Epidermal growth factor (EGF
ESRP1	Epithelial splicing regulatory proteins 1

ESRP2	Epithelial splicing regulatory proteins 2
EPB41L3	Erythrocyte Membrane Protein Band 4.1 Like 3
ES	Exon skipping
ESEs	Exonic splicing enhancers
ECM	Extracellular matrix
FDR	False discovery rate
FGF	Fibroblast growth factor
FIT	Fecal immunochemical test
FTMS	Fourier transform mass spectrometry
FOBT	Fecal occult blood test
FOXP1	Forkhead Box P1
FTMS	Fourier transform mass spectrometry
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
gFOBT	The guaiac fecal occult blood test
GSK-3β	Glycogen synthase kinase 3-β
HARS	Histidyl-TRNA Synthetase
hnRNPs	Heterogeneous nuclear ribonucleoproteins
HPRT1	Hypoxanthine phosphoribosyltransferase1
HDI	Human Development Index
HOGG1	Human oxoguanine glycosylase 1
ELISA	Immunosorbent assay
IGFBP-2	Insulin-like growth factor binding protein 2
IGF	Insulin-like growth factor
ILK	Integrin-linked kinase
lncRNAs	Long noncoding RNAs
MALDI	Matrix-assisted laser-desorption ionization
MAPKs	Mitogen-activated protein kinases
MASP1	Mannan binding lectin serine protease 1
MEF	Mouse embryonic fibroblast

miR-140	MicroRNA 140
MMPs	Matrix metalloproteinases
MS	Mass spectrometry
MBNL	Muscleblind-like
MDS	Myelodysplastic syndromes
M	Metastasis
miRNAs	MicroRNAs
MMR	Mismatch repair
MSI	Microsatellite instability
MSS	Microsatellite instability
mRNA	Messenger RNA
NCBI	National Center for Biotechnology Information
N	Node
NI	New intron
NLR	The neutrophil-to-lymphocyte ratio
OS	Overall survival
PCR	Polymerase chain reaction
PI3Ks	Phosphatidylinositol 3-kinases
PPT	Polypyrimidine tract
qPCR	Quantitative real-time PCR
RBPs	RNA-binding proteins
RE	Retained exon
RefSeq	Reference Sequence Database
RFS	Recurrence-free survival
RNA	Ribonucleic acid
RBP	RNA binding proteins
RNP	RNA-protein
RNAseq	RNA sequencing
RT	Reverse Transcriptase

snRNP	Small ribonucleoprotein
shRNA	Small hairpin RNA
SATB2	Special AT-rich sequence binding protein 2
SRE	Splicing regulatory elements
SR	Serine–Arginine
STK4	Serine/threonine kinase 4
TGF	transforming growth factor
T	tumor
TP53	tumor protein p53
TSGs	tumor suppressor genes
UICC	Union for International Cancer Control
VEGF	vascular endothelial growth factor
wt	wild type
ZnF	Zinc finger
ZFAS1	ZNFX1 antisense RNA1
3′SS	3′ splice site
5′SS	5′ splice site

2 Introduction

Despite the latest progress in methods of screening and diagnosis, colorectal cancer (CRC) is still one of the most important leading causes of cancer related death worldwide, and it remains the third most commonly diagnosed malignancy in males and the second most commonly diagnosed cancer in females [1]. Many factors including somatic mutations in KRAS, adenomatous polyposis coli (APC), Epidermal growth factor receptor (EGFR) and TP53 serve as positive prognostic markers in patients with colorectal tumor progression [2]. The incidence of CRC in inhabitants less than 50 years of age (early onset disease) has seen an alarming rise worldwide over the past two decades. Furthermore, early-onset CRC administer to offer with higher pathologic grade, distant disease, and a greater incidence of recurrence and metastatic disease. The reason for this global trend is poorly understood [3]. However, currently established prognostic and predictive markers are not sufficient for a next dramatic change in the CRC outcome. Novel markers are needed, and recent studies add growing evidence of aberrant alternative splicing (AS) importance as prognostic, diagnostic, and predictive biomarker in development and progression of colorectal cancer, metastasis, therapeutic resistance, and other oncogenic processes [4]. Alternative splicing may bring a promise of faster diagnosis and prognosis in cancer field in general, not only in CRC [5]. Alternative splicing is efficiently regulated by splicing regulatory factors such as muscle blind-like (MBNL) protein family. In addition to directly change effector proteins, any mutations or changes in expression of splicing factors themselves could lead to the activation of oncogenes and cancer pathways or the loss-of-function in tumor suppressors [4].

Figure 1. Anatomy of the colon,Colon cancer and polyp
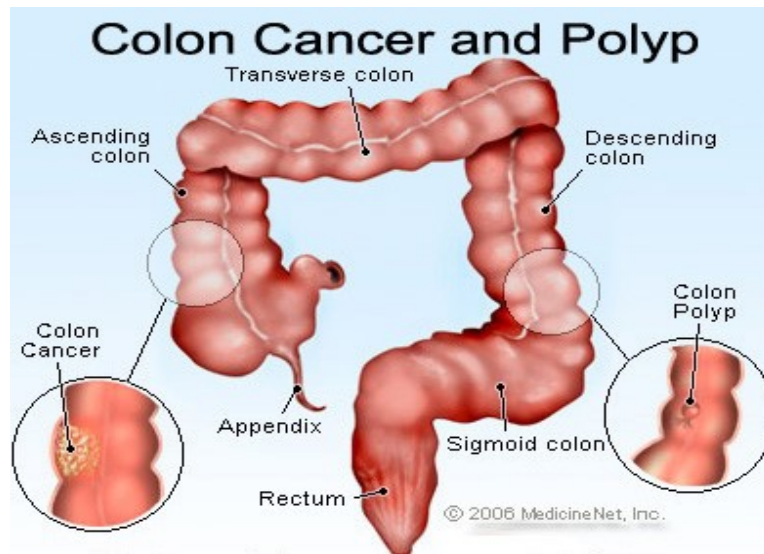
Figure 1. The human colon is arbitrarily divided into four sections, including ascending, transverse, descending, and sigmoid colon. Benign polyps develop when DNA damage occurs in the inner lining of the colon cell (Right). Over the years, polyps in the colon are able to develop into colon cancer (Left). Illustration from Medicine Net. Inc

https://www.medicinenet.com/image-collection/colon_cancer_picture/picture [6].

2.1 Introduction to Colorectal Cancer
2.1.1 CRC Epidemiology

Colorectal cancer (CRC), also known as colorectal adenocarcinoma, generally originate from the glandular and epithelial cells of the large intestine when some of the epithelium cells undergo a series of genetic changes that transform them into adenocarcinomas [7] (Figure 1). Along with the abnormal accumulation of epigenetic changes in colon epithelial cells, these hyper-proliferative cells generate rapidly growing benign adenoma, which may later develop into carcinoma and over decades into development of metastases [8,9].

According to GLOBOCAN 2020 data, CRC is the third most common cause of cancer related death in the world with about 1 million deaths per year [10]. Almost 2 million new cases and about 1 million deaths are expected to happen yearly in the next 10 years [1,11]. Nevertheless, it has consistently been shown that due to population growth and ageing, CRC incidence and the number of cancer deaths is increasing

rapidly in the entire world (Figure 2), especially in medium to high Human Development Index (HDI) countries particularly in Eastern Europe, Asia and South America that are undergoing rapid transition in societal and economic and adopting western way of life associated with using up more red meat, obesity, sedentary lifestyle, alcohol and tobacco consumption, that are mostly considered as the driving factors behind the growth of CRC [8,12]. According the latest data, CRC accounts for the deadliest cancer among males in three countries - Saudi Arabia, Oman, and UAE, and the deadliest among females in five countries including Algeria, Belarus, Japan, Spain, and Portugal [13]. The country with the highest mortality rates of CRC per 100,000 population is Hungary among both males and females [11,14]. According to GLOBOCAN 2020 data, total CRC incidence and mortality rates have been significantly decreasing in a number of the highest indexed HDI countries, for example in the USA, Australia, New Zealand and western European countries. In contrast to declining of CRC mortality in these countries, however, the incidence of CRC in inhabitants less than 50 years of age (early-onset disease) has been increasing. Early-onset CRC incidence in the U.S. is going up by an average of 1.8% annually from 1992–2012, and is estimated to account for 10% to 25% of newly-diagnosed CRC by 2030 [3]. In 2020, CRC is reported for 10% of global cancer incidence and 9.4% of cancer deaths, quite lower than lung cancer that represents 18% of deaths. The global number of new CRC cases is predicted to reach 3.2 million in 2040, based on the prediction of population growth, aging, and human development. The increase in CRC incidence is mainly ascribed to the increase exposure to environmental risk factors resulting from lifestyle changes and diet toward western way of life. The International Agency for Research on Cancer (IARC) evaluates that the global burden of CRC will grow by 56% between 2020 and 2040, to more than 3 million new cases per year [15]. The predicted increase in the number of deaths from the disease is even more, by 69%, to about 1.6 million deaths worldwide in 2040 [10]. Number of predicted deaths and mortality rate for the year 2020 in comparison with 2015 displays Table1. There are approximately 450,000 people newly diagnosed with CRC each year in Europe [11].

Colorectal cancer predicted mortality rates for 2020 were 15.8 per 100 000 men (−6.7%) and 9.2 in women (−7.5%) and their declines have been reported in all age

groups [12]. As it is presented in Figure 3, the averaged crude mortality rate between 2015 and 2020 has been found substantially reduced by 5% in men and 4% in women. The reasons for the recent declining trends in mortality in these countries may be partially explained by improvements in early detection, diagnostic techniques, screening programs and prevention through polypectomy [6].

Number of mortality and death rates for year 2020

Table 1. Number of predicted deaths and mortality rate for the year 2020, comparison figures for 2015, for the EU, with 95% prediction intervals (PIs) and the percentage differences between 2015 and 2020					
	Observed number of deaths 2015	Predicted number of deaths 2020 (95% PI)	Observed ASR 2015	Predicted ASR 2020 (95% PI)	% difference 2020/2015
Men					
Stomach	34 666	33 700 (32 971—34 485)	6.28	5.53 (5.39—5.67)	−11.9
Colorectum	93 241	98 500 (96 750—100 316)	16.06	15.38 (15.08—15.68)	−4.2
Pancreas	42 462	45 900 (45 350—46 534)	7.91	7.76 (7.65—7.87)	−1.9
Lung	183 943	182 600 (179 799—185 440)	34.83	31.62 (31.02—32.21)	−9.2
Prostate	74 998	78 800 (77 646—80 026)	10.71	9.95 (9.78—10.11)	−7.1
Bladder	31 938	32 800 (32 036—33 618)	4.99	4.52 (4.38—4.65)	−9.4
Leukemias	23 713	25 100 (24 489—25 707)	4.39	3.87 (3.73—4.02)	−11.8
All cancers (malignant and benign)	760 111	798 700 (791 039—806 298)	137.53	130.09 (128.63—131.55)	−5.4
Women					
Stomach	22 162	20 300 (19 698—20 900)	2.91	2.51 (2.42—2.6)	−13.7
Colorectum	77 122	76 100 (74 944—77 260)	9.40	8.62 (8.48—8.75)	−8.3
Pancreas	42 201	46 200 (45 494—46 975)	5.53	5.60 (5.49—5.7)	1.2
Lung	88 502	99 800 (98 658—100 983)	14.28	15.13 (14.95—15.32)	6.0
Breast	93 903	95 900 (94 637—97 187)	14.54	13.48 (13.31—13.66)	−7.3
Uterus (cervix and corpus)	29 691	31 300 (30 656—31 974)	4.84	4.78 (4.67—4.89)	−1.2
Ovary	30 213	29 000 (28 330—29 696)	4.78	4.23 (4.11—4.36)	−11.5
Bladder	10 900	11 400 (11 114—11 656)	1.17	1.15 (1.12—1.18)	−1.9
Leukemias	19 137	19 800 (19 390—20 184)	2.63	2.27 (2.17—2.37)	−13.7
All cancers (malignant and benign)	603 957	630 100 (625 252—634 898)	85.68	82.16 (81.59—82.73)	−4.1

ASR, age-standardised mortality rates using the world standard population.

Table 1, the averaged crude mortality rate between 2015 and 2020 has been found substantially reduced in both men and female https://doi.org/10.1016/j.annonc.2020.02.009[16].

Figure 2. number of new cases in 2020, worldwide, all ages
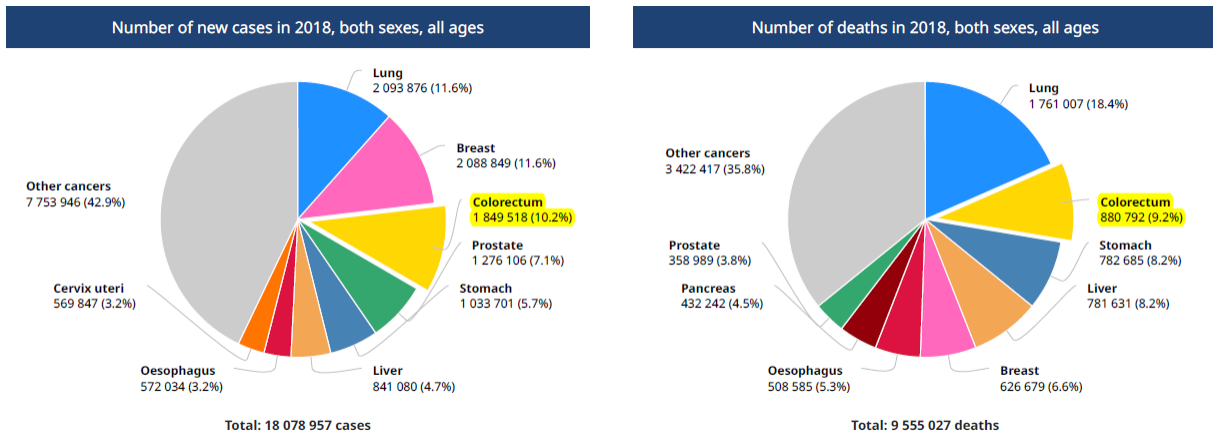
Figure 2. Global colorectal cancer incidence and mortality of all ages in the world.

Source: GLOBOCAN 2020: New Global Cancer Data | UICC https://www.uicc.org/news/globocan-2020-new-global-cancer-data [17]

Figure 3. EU age-standardised mortality rates observed for 2015 and predicted for 2020 with corresponding 95% PIs for total cancers and the 10 major cancer sites in men and women
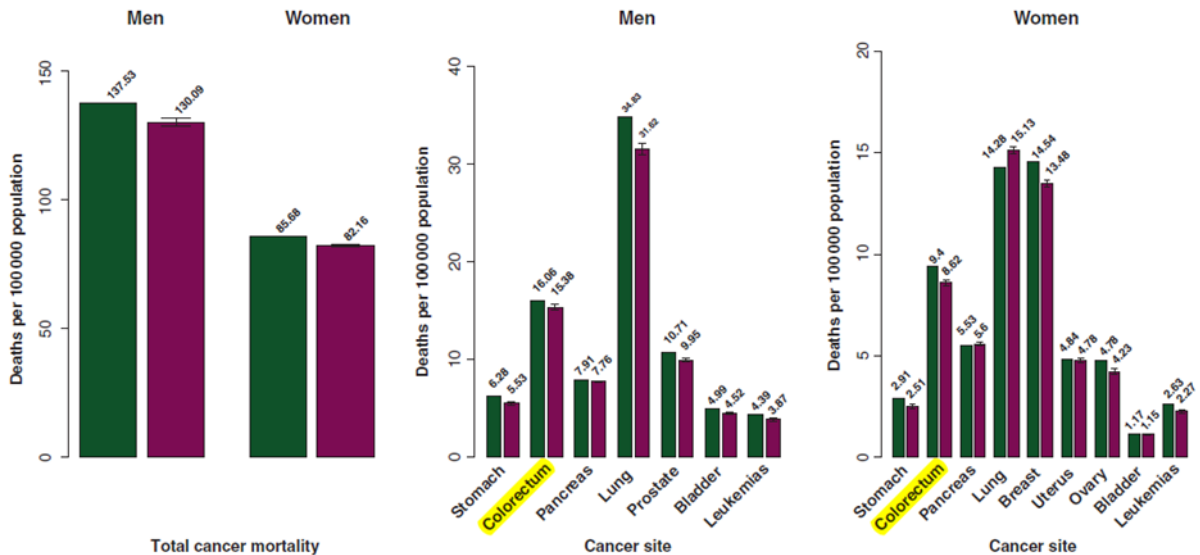
Figure 3. Bar-plots of age-standardized death rates per 100,000 people a year in 2015 and predicted rates in 2020 for all cancers and 10 major cancer sites in EU men and women. Source: Annals of Oncology https://doi.org/10.1016/j.annonc.2020.02.009[16]

2.1.2 Biomarkers in CRC

The ideal molecular marker can be defined as a substance that can be quickly, reliably, reproducibly and quantitatively measured, with high sensitivity and specificity, that is consistently detected across genders and ethnic groups while its concentration is reflecting the stage of the disease and/or the response to the therapy used [18]. Thus protein biomarkers are considered as a robust indicators of physiological biological processes, pathogenic processes, or pharmacologic responses in establishment of effective treatment strategy [19]. In addition, peripheral blood biomarkers that illustrate an individual's risk for developing CRC or other cancers, risk related Single nucleotide polymorphisms (SNPs) and the determination of DNA methylation in white blood cell DNA as a marker of cancer risk as well as other molecular markers can be used to assess the risk of future disease, the aggressiveness of the malignancy over the time, and the probability that a patient will respond to a particular treatment, thereby helping the clinician to make personalized treatment decisions. Recently, translational research in CRC has produced a wide spectrum of potential molecular biomarkers in CRC tumor tissues which illustrate the clinical applications of biomarkers in early detection of the disease and their essential role in prognostic stratification, diagnosis, surveillance, and treatment of CRC [20]. However, it must be taken into account that due to high heterogeneity of CRC and poor prognosis associated with late diagnosis, there is an essential need for new diagnostic, prognostic and predictive biomarkers to avoid CRC related deaths [21]. CRC is specific by its high inter and intra tumor heterogeneity with various mutational landscapes, that arise via the genomic microsatellite instability (MSI), chromosomal instability (CIN), aberrant DNA methylation, and DNA repair defects [22]. Most CRC cases are due to sporadic genetic or epigenetic changes, with up to 10% to 20% of all CRC cases have a familial history of the disease [22] and only 2% [23,24] of all CRC cases have described germline mutations associated with hereditary cancer syndromes [15]. It must be considered that genetic and epigenetic changes characterizing the carcinogenesis of CRC are essential for the identification and development of a suitable biomarker [23]. Sporadic colorectal carcinogenesis is the term given to complex multifactorial processes resulting in the alteration of normal epithelial cell cycle that occurs in individual cases with no known familial history of a hereditary CRC syndrome [24]. The acquisition of genomic instability is considered a key hallmark in tumor development. Based on the

major driving mechanisms behind its development, they can be characterized by chromosomal instability (CIN), which is the most common pathway to sporadic CRC (Table 2), the CpG island methylator phenotype (CIMP), microsatellite instability (MSI), and a KRAS-mutated and BRAF wild type (wt) gene [25,26]. The CIN, MSI and CIMP pathways result in lesion pathology and progression transition to malignancy, which is going along with deregulated gene expression of oncogenes and tumor suppressor genes. CIN is the most common subset of genomic instability, responsible approximately for 80%–85% of colorectal tumors. This subset shows different forms of genomic instability, specified mainly by chromosomal rearrangements and numerical abnormalities at a large increased rate compared with normal cells. At the outset, the CIN phenotype employed to characterize tumors with a high degree of intercellular heterogeneity in chromosome number, specified by counts for a restricted set of chromosome-specific centromeres. CIN is further used to hallmark cancers with either imbalance in chromosome number (aneuploidy) or polyploid DNA content as measured by cytometry or cytogenetics, or multiple gains or deletions of chromosomes or chromosome arms, high frequency of loss of heterozygosity (LOH). MSI CRC occurs in nearly 7%–12% of sporadic colorectal cancers. This subtype results from an impair of DNA mismatch repair function or partial MMR inhibition, for example by incomplete hypermethylation of MMR genes or due to a production of alternative variants of MMR genes [25].

Another pathway of tumorigenesis, the CpG Island Methylator Phenotype (CIMP) that also occur through an epigenetic instability pathway, is characterized by vast hypermethylation of promoter CpG island sites, resulting in the inactivation of several (at least 5) tumor suppressor genes or other tumor-related genes. The molecular features of these pathways can potentially be used clinically in the diagnosis, screening and management of patients and families with colorectal cancer. On that account, the substantial genetic heterogeneity in colorectal tumors must be considered when developing novel molecular diagnostic methods [27].

Recent advances in molecular and biological characteristics of CRC have provided useful knowledge about its pathogenesis. It leads to the development of biomarkers that help with the identification of patient responses with respect to cancer diagnosis, management, and surveillance. The detection of biomarkers that can result in CRC early detection or monitoring would enable the development of proper prevention,

personalized medicine and improve survival rates [26]. Cancer biomarkers could be divided according to their different functions on diagnostic, predictive, or prognostic [27].

Diagnostic biomarkers are known as a biological characteristic that can be used in early detection of the cancer in a noninvasive way and thus have the secondary function for a prevention of the cancer [28,29]. Prognostic markers provide prognostic information, independent on therapy, on patient outcomes at the time of diagnosis, such as overall survival (OS) and relapse-free survival (RFS). In contrast, predictive biomarkers are usually used to provide information about treatment response and therefore could be used to guide therapy decisions [20,30].

Selected Genes That Regulate Chromosomal Instability

Pathway	Deregulated gene	Gene alteration
Chromosome segregation	*BUB1; BUBR1*	Somatic mutations
	hZw10; hZwilch; hRod	Somatic mutation
	AURKA	Gene amplification and overexpression
	PLK1	Gene overexpression
	PTTG	Gene overexpression
	CENP-A; CENP-H; HEC1; INCENP	Gene overexpression
	APC	Germline and somatic mutations
Telomere regulation	*TERC*	Overexpression
DNA damage response	*ATM; ATR*	Germline mutation
	BRCA 1/2	Germline mutation
	TP53	Germline mutation and somatic mutations
	MRE11	Germline mutation

Table 2. Responsible Genes that regulate chromosomal instability https://www.slideshare.net/spaliz/a-the-chromosomal-instability-pathway-in-colon-cancer-184158841[31]

2.1.3 Importance of Biomarkers in CRC

It is important to detect new CRC biomarkers with high specificity, sensitivity and reproducibility. Despite the fact that in some cases these putative markers appear as good diagnostic parameters, in general it has been shown that a single protein marker is not enough to get high sensitivity and specificity values for CRC diagnosis [30,32]. Therefore, employing a panel of several biomarkers might be useful to achieve a greater diagnostic accuracy.

In addition to the diagnostic biomarkers, clinical management of CRC patients can further benefit from implementation of variety of prognostic and predictive biomarkers. A prognostic biomarker is a clinical or biological characteristic which is specified to a special tumor type by defining the occurring polymorphism, mutation or the change in DNA methylation or gene expression, or by detecting the presence of specific microRNA (miRNA) molecules or circulating tumor cell (CTCs) in the peripheral blood and provide information on patient health outcome regardless the treatment, while a predictive biomarker presents information on the possibility of obtaining a response to treatment and indicates the likely benefit the patient get from a specific treatment, compared to their condition at baseline [30,33]. The most important predictive biomarkers with a recognized clinical value in the molecular diagnostics in oncology are somatic mutations, of both the point mutation and chromosomal aberration types, in the following genes: EGFR, KRAS, BRAF, PDGFRA, KIT, HER2, BCR-ABL, and EML4-ALK [34,35].

In the following chapters, we review the present status of newer diagnostic, prognostic, and predictive biomarkers in CRC and provide perception of their implementation in clinical management. All biomarker types will be divided according to sample origin into Tissue biomarkers, Blood Biomarkers and Stool Biomarkers [36].

2.1.4 Diagnostic biomarker

As mentioned above, diagnostic biomarkers can be identifying the presence of a disease or condition of interest, in some cases also a specific subtype of the disease, optimally prior to occurrence of symptoms, allowing for early detection of the disease. Latest discoveries have been demonstrating developments of *CRC* screening modalities including invasive and non-invasive biomarkers. Both minimally-invasive and non-invasive biomarkers are required to encourage use by clinicians and increase

compliance of patients (Table 3). To obtain these properties, biomarkers in blood, urine, saliva, feces or pancreatic juice samples must be investigated.

1. Tissue Biomarkers

1) Cytokeratins

Cytokeratins (CKs) are intermediate filaments forming proteins found in the intracytoplasmic cytoskeleton of epithelial tissue, provide mechanical support and fulfil a variety of additional functions in epithelial cells [37]. CK staining models are the most helpful procedures for discovery of metastatic adenocarcinoma of unknown primary origin [38]. To differentiate metastatic CRC from other tumors, tissues are stained for CK7 and CK20. CRC specimens mostly stain positive for CK20 and negative for CK7 [39,40]. CK20 selectively exist in Merkel cells and normal gland cells of the colonic mucosa, while CK7 is expressed in urinary bladder and female genital tract epithelia, mesothelium, and normal lung and is not present in the colonic mucosa. Detecting the CK7–/CK20+ pattern is a typical method for metastatic CRC diagnosis. A reported 65% to 95% of CRC cases show a CK7–/CK20+ pattern [41].

2) Caudal type homeobox 2

Caudal type homeobox 2 (CDX2) codes for a homeobox protein which is involved in the regulation of normal cell differentiation in the GI tract and tumor suppression in the colon [42]. While CRC is frequently associated with abundant CDX2 expression, loss of CDX2 is linked with adverse prognosis. Werling et al. highlighted tumor-suppressor role of CDX2 expression in CRCs. Their study showed the correlation between loss of CDX2 expression in CRCs with tumor differentiation, proximal tumor location, microsatellite instability (MSI), CpG island methylator phenotype (CIMP) and BRAF mutation [43].

Many studies have evaluated that high CDX2 expression levels were found in CRC, while intermediate CDX2 expression levels were found in other adenocarcinomas of the GI tract. Other than CK pattern, assessing CDX2 expression is also a very sensitive and specific manner of identifying CRC [44].

3) Special AT-rich sequence binding protein 2 (SATB2)

Special AT-rich sequence binding protein 2 (SATB2) is known as part of the nuclear matrix attachment region DNA binding transcription factors family [45]. Although the exact role of SATB2 in the GI tract is unknown, Magnusson et al. found that SATB2 being highly expressed in the epithelium of the lower GI tract including the appendix, colon, and rectum [46]. SATB2 could be used as a single marker of CRC with shown positivity of 87.8% (943/1,074). Many studies have reported the benefit of the addition of SATB2 to the standard panels of CK7, CK20, and CDX2, however no clear consensus has been reached [47]. Dragomir et al. described that the addition of SATB2 to standard panels has no significant improvement in sensitivity or specificity in the diagnosis of CRC, which is in contrast with recent studies reporting that SATB2 might be a specific marker for differentiating metastatic CRC from primary ovarian carcinomas [48]. Before the identification of SATB2, CDX2 was the most specific marker available [49]. However, CDX2 tested positive in 18% of ovarian carcinomas, whereas SATB2 was expressed in 95% of metastatic CRC and 0% of ovarian carcinomas. Therefore, SATB2 could be used as a marker to exclude ovarian carcinomas [50].

4) Cadherin 17

Cadherins are cell–cell adhesion molecules and have critical roles to keep proper tissue structure under physiological conditions. Molecular defects in cadherin expression are involved in many human diseases including carcinomas [51]. Cadherin 17 (CDH17), a member of the cadherin superfamily, was first identified in the rat liver and intestine. Later on, 3 immunohistochemical studies consisting of a large number of human tissues found that CDH17 was expressed on the surface epithelium of the duodenum, ileum, appendix, and colon [52]. Recent studies illustrated the role of CDH17 as immunohistochemical marker for the diagnosis of adenocarcinomas of the GI tract [53,54]. CDH17 is reportedly expressed in 96% to 100% of primary and 100% of metastatic CRC [55,56]. CDH17 is also expressed in other GI tumors such as gastric, pancreatic, and biliary cancer but it is rarely found outside the GI tract. Although CDH17 is transcriptionally regulated by CDX2, some studies indicated that CDH17 is more sensitive and specific than CDX2 for the identification of CRC [51].

5) Telomerase

Telomerase, also called terminal transferase, is a ribonucleoprotein that maintains telomeres by adding TTAGGG repeats onto telomeres that are located at the ends of chromosomes [57,58]. Telomerase uses intrinsic RNA as a template for reverse transcription. Under physiological conditions, telomeres shorten with each cell division. When telomeres become critically shortened, a signal of DNA damage is induced that results in replicative senescence [58]. Cancer cells bypass DNA damage-induced inhibitory signaling pathways by upregulating telomerase. Telomerase is found in 85% to 90% of all malignant tumors [59]. The activity level of telomerase reflects dysplasia progression in the respective tissue. Telomerase can be introduced as a novel diagnostic biomarker in CRC using the new methods of Telomerase Biosensor Technology [60]. For example, Heidenreich and Kumar discussed the importance of telomerase in human cancers, described the mechanisms of telomerase rehabilitation with focus on cancer specific noncoding mutations in the core promoter of the TERT gene that result in creating of new binding sites for E-twenty-six (ETS) transcription factors. The telomerase reverse transcriptase (TERT) promoter mutations marked the first case of driver mutations within the so named dark matter of the genome which act through aberrant expression rather than alteration of a gene products[61].

6) GPA33

The GPA33 (A33) gene encodes protein A33, a membrane-bound glycoprotein homologous to the immunoglobulin superfamily [62,63]. Immunohistochemical studies have found that A33 is expressed by epithelial cells in the stomach, small intestine, colon, and rectum. It is expressed in more than 95% of human CRC cases, especially in well-differentiated tumors, informing that it could be a potential target for CRC treatment [56]. An immuno-histological study comparing A33 and CDX2 determined that A33 showed sensitivity similar to that of CDX2 but specificity higher than that of CDX2 as an immunomarker of CRC [64].

2. Blood Biomarkers

1) Chromosomal instability associated alterations

Chromosomal instability (CIN) is known as gain or loss of whole or large portions of chromosomes leading to karyotypic variability, resulting in sub-chromosomal genomic amplifications, changes in chromosome numbers, and a high loss of heterozygosity (LOH) rate [65,66]. CIN is the most common form of genetic instability observed in CRC (65%–70%) [67]. Mutations in tumor suppressor genes (TSGs) and oncogenes such as APC, CTNNB1, KRAS, PIK3CA, and TP53, and LOH of chromosome 18q were the key events that lead to the development of CIN CRC [68,69]. Various cases of colorectal cancer result through CIN pathway, which is typical by general imbalances in chromosome number (aneuploidy) and loss of heterozygosity (LOH) [65,66]. Interestingly, it was recently shown that CIN may occur as early as in low grade dysplasia colon adenomas [70]. The CIN that underlies most of all colon cancers, can induces carcinoma through the loss or mutation of tumor suppressor genes such as APC, TP53 and also through activation of oncogenes such as RAS family. RAS family includes HRAS, NRAS and KRAS genes. Among RAS family, KRAS gene plays the most important role. KRAS gene encodes GTP (guanosine 5′-triphosphate)-binding proteins, and they function as self-inactivating intracellular signal transducers for surface receptors such as epidermal growth factor receptor. The activation of RAS genes results in cell survival promotion and suppress apoptosis[69]. As well as KRAS, v-raf murine sarcoma viral oncogene homologue B1 (BRAF) genes encode serine/threonine protein kinase that is regulated by KRAS protein [68]. It must be also considered of Adenomatous polyposis coli (APC) gene role in the Wnt/Wingless pathway which makes this gene as critical gatekeeper of colonic epithelial cell proliferation. APC gene is responsible for controlling the underlying oncoprotein called β-catenin. The loss of function in APC gene results in transition of adenoma from normal colonic mucosa due to the up-regulation of β-catenin. Blood sample analysis can detect presence of CRC-related germ-line mutations. Germline APC mutations mostly give rise to familial adenomatous polyposis (FAP) which is an inherited cancer-associated derangement. Such as KRAS, APC mutations occurs in the early stage of the progression from adenoma to carcinoma. Therefore, mutations in the KRAS and APC gene are good biomarkers and can be used to identify individuals associated with

high risk of CRC in patients' families so is used as guidelines for the frequency of CRC surveillance and the recommendation of prophylactic surgery [71].

2) Circulating cell-free DNA

Circulating cell-free DNA (cfDNA) is defined as a type of cell-free nucleic acid that is released from apoptotic or necrotic cells into blood circulation [31,72,73].The lengths of cfDNA strands differ among processes by which they are made. In healthy individuals, cfDNA is released from apoptotic cells, and the DNA fragments are about 180 bp long [74,75]. However, in tumor cells, cfDNA is released as much larger fragments by necrosis. Quantitative analysis of circulating cfDNA in CRC patients and healthy controls was associated with relatively low sensitivity (73.5%) but acceptable specificity (91.8%) for CRC [56]. Moreover, integrity index reflecting ratio of longer and shorter DNA fragments displays better diagnostic accuracy compared to absolute cfDNA concentrations.

In addition to cfDNA only, methylation status of cfDNA can be used as a diagnostics marker, although currently it appears to be less specific for CRC compared to circulating tumor DNA (ctDNA) only.

Although methylated cfDNA represents one of the most promising candidates for blood based markers with good sensitivity for early-stage CRC, further research is required to find which methylated cfDNA markers are the most accurate when applied to large cohorts of patients.

3) MicroRNA

MicroRNAs (miRNAs) are 18–25 bp-long small noncoding RNAs that regulate gene expression by binding to mRNA [76]. The miRNAs are reportedly correlated with different cancers including CRC by acting as oncogenes or tumor suppressor genes (TSGs). Compared to mRNA, miRNAs show higher stability in the blood since they avoid degradation by endogenous RNase and are resistant to extreme pH changes [77]. For these reasons, miRNA is likely-looking noninvasive biomarkers in cancer [48]. Studies have distinguished their diagnostic ability for CRC using both single miRNA as well as specific miRNA panels [78].

Many studies illustrated that main serum miRNAs such as miR-92, miR-141, miR-29a, miR-21, miR-221, miR-601 and miR-760, might be used as a promising diagnostic marker for CRC. For example, some researchers reported that miR-21 showed 90% specificity and sensitivity for CRC detection [79,80].

4) Long noncoding RNAs

Long noncoding RNAs (lncRNAs), that consist of more than 200 nucleotides but are not coding for a protein, are involved in various biological processes such as epigenetic regulation, immune responses, cell differentiation, and chromosome dynamics [81]. More than 150 human diseases were linked to lncRNAs, including colon cancer, breast cancer, leukemia, and psoriasis. Studies have highlighted their diagnostic ability for CRC using single or panel lncRNAs [82,83]. According to the targeting regulatory mechanisms of lncRNAs, these potential biomarkers provide following four functions: signal function which regulate the spatiotemporal expression of target genes, scaffold function which act as adapters in functional protein complexes, guide function which act by binding to specific proteins and direct the localization of the resulted complex, and bait function which prevent other RNAs or proteins from binding to their natural targets [84]. lncRNAs can act as effectors to direct biological processes. Namely, there is a considerable number of lncRNAs such as NEAT1, ZNFX1 antisense RNA1 (ZFAS1), and GAS5 with hopeful potential function which participate in both normal and aggressive disease state as CRC diagnostic or prognostic biomarkers [85]. For example, it was demonstrated that CRC patients have higher levels of long non-coding Homeobox Transcript Antisense Intergenic RNA in the blood compared to healthy controls. Moreover, high level of its expression in blood and primary tumor were associated with worse prognosis in these patients [86].

5) Insulin-like growth factor binding protein 2

Insulin-like growth factor binding protein 2 (IGFBP-2) is a binding protein that modulates the interaction between insulin-like growth factor (IGF) ligands and IGF-1 receptors[87]. Despite of unknown physiological role of IGFBP-2, several studies have assessed that elevated serum levels of IGFBP-2 are associated with malignancies of the colon, ovary, and prostate as well as other advanced solid tumors [88]. Liou et al.

found that elevated serum and plasma IGFBP-2 levels could distinguish patients with CRC or colon polyps from healthy controls [88,89]. Although the sensitivity and specificity of IGFBP-2 alone is unsatisfactory for early CRC and colon polyp detection, its combination with other biomarkers such as CEA could increase the sensitivity [90]. Also, higher plasma IGFBP-2 levels were associated with larger tumor size and worse overall survival (OS) rates, indicating that IGFBP-2 might serve as a diagnostic and prognostic biomarker for CRC [91].

3. Stool Biomarkers

Stool-based assays are considered the most useful type for many reasons. According to numerous authors, CRC and polyps exfoliate many neoplastic cells and their debris into the mucocellular layer of the colonic lumen [92]. The detectable molecular changes that are caused by CRC cells are reportedly present in the stool earlier than in the blood [93]. The actual unfolded surface area of the epithelial monolayer of cancers and polyps could be 200 times larger than that predicted by gross findings [94]. Therefore, the detection of components of the exfoliated cells, such as DNA, miRNAs, and proteins, might represent a potent approach in CRC diagnosis.

1) Fecal occult blood and immunochemical test

The guaiac fecal occult blood test (gFOBT) and fecal immunochemical test (FIT) are known as common non-invasive techniques for screening for CRC [95,96]. Based on randomized control clinical trial reports, gFOBT as a CRC diagnostic test led to mortality reduction by 11%–33% over the last 20 years [97]. Although it can be used as a biomarker, it could not recognize origin of bleeding between the upper and lower GI tract. Accordingly, it shows only 30% to 40% sensitivity for CRC detection. In detecting adenomas over 1 cm in diameter, FIT showed only 20% to 30% sensitivity [98,99]. Furthermore, the occult blood test can detect left-sided lesions in the colon much more than right-sided lesions [100–102]. This is considered as most severe limitation since the incidence of right-sided CRC has been increasing during last decades. Therefore, neither gFOBT nor FIT can be considered as independent diagnostic tools [103,104].

2) Stool DNA

The stool DNA test, also called a stool DNA-fecal immunochemical test, is a noninvasive laboratory test that identifies DNA changes in the cells of a stool sample. The stool DNA test is a novel method to screen for diagnosing CRC [105,106]. Less than 0.01% of the total DNA in the stool is human DNA, while the other 99.99% is derived from intestinal bacteria or from the diet [56]. For diagnosis of the CRC, not only the presence of particular sequence is important, but also its modification [107–109]. Analysis is mostly focused on methylation status of detected DNA, and the genes tested are for example APC, ATM, BMP3, CDH1, KRAS, CDKN2A, CDH13, VIM, and WIF1, which have been analyzed in fecal DNA for the early detection of CRC [109–111].

3) Stool miRNA

MicroRNA (miRNA) patterns in colorectal tissues could be reproduced in fecal samples. Since the environment of the GI tract is much more complicated than that of the blood, marker stability within it is a major concern [112,113]. Many studies have illustrated that, unlike rapid degradation of mRNA and protein, miRNA transcripts were more stable in different conditions. Although there are less reports about miRNA in the stool compared to those in the blood, dysregulation of miRNA expression was discovered in the stool of CRC patients [114]. Saray Duran-Sanchon et al. have assessed that miR-92a, miR-21, miR144, miR-106a, miR17-92 cluster, and miR135 had increased expression in CRC while miR-143 and miR-145 were downregulated in CRC [115]. However, none of the miRNAs showed adequate predictive value for use as a CRC diagnostic test alone, and more information are needed to increase the diagnostic value of miRNA by combining several miRNAs into the detection panel [116,117].

Summary of Diagnostic Biomarkers

Biomarker	Explanation
Tissue biomarker	
Cytokeratins (CKs)	65%-95% of CRC cases show a CK7-/CK20+ pattern.
Caudal type homeobox 2 (CDX2)	Sensitivity and specificity of CDX2 expression in CRC diagnosing is greater than 90%. As CDX2 expression alone cannot differentiate among adenocarcinomas of the GI tract, [13,14] it is useful as an adjunct to CKs.
Special AT-rich sequence binding protein2 (SATB2)	The addition of SATB2 to standard panels showed no significant improvement in sensitivity or specificity in the diagnosis of CRC. As SATB2 expression was positive in 95% of metastatic CRC and 0% of ovarian carcinoma cases, it could be useful as a marker to exclude ovarian carcinomas.
Cadherin 17 (CDH17)	Cadherins are cell-cell adhesion molecules that play important roles in maintaining tissue structure under normal conditions. CDH17 is reportedly expressed in 96%-100% of primary and 100% of metastatic CRC.
Telomerase	Telomerase is a ribonucleoprotein that maintains telomeres by adding TTAGGG repeats onto them. Cancer cells bypass DNA damage-induced inhibitory signaling pathways by upregulating telomerase. Using TBT, telomerase showed 95% sensitivity and 95% specificity in CRC.
GPA33 (A33)	A33 is expressed in the stomach, small intestine, colon, and rectal epithelial cells. It is expressed in more than 95% of human CRC. Also, an immunohistological study comparing A33 and CDX2 revealed that A33 showed similar sensitivity as but a higher specificity than CDX2 as an immunomarker of CRC.
Blood biomarker	
Circulating cell-free DNA (cfDNA)	The cfDNA is released as much larger fragments in tumor cells than in normal cells. Quantitative analysis of circulating cfDNAs by measuring the ratio of longer and shorter DNA fragments, in other words, measuring cfDNA integrity number during the diagnosis of CRC, showed sensitivity of 73%-90% and specificity of 97%-85%.
MicroRNA (miRNA)	The miRNAs are 18-25 bp-long small noncoding RNAs that regulate gene expression by binding to mRNA. Compared to mRNA, miRNAs show higher stability in the blood. Different combinations of miRNA showed high sensitivity and specificity for detecting CRC.
Long noncoding RNA (lncRNA)	HIF1A-AS1 showed high diagnostic ability of CRC with an AUC of 0.960. CRC patients with high HIF1A-AS1 expressions were associated with shorter 5-year survival rate than those with low HIF1A-AS1 expression.
	CRNDE-h showed better diagnostic value compared to CEA. When combined with CEA, he diagnostic value improved.
	Other markers such as NEAT1, ZFAS1, and GAS5 showed promising results for potential use as a diagnostic or prognostic marker.
Insulin-like growth factor binding protein 2(IGFBP-2)	An elevated serum IGFBP-2 level is associated with malignancies of the colon, ovary, and prostate as well as other advanced solid tumors. [62-65] Although the sensitivity and specificity of IGFBP-2 alone is unsatisfactory for early CRC and colon polyp detection, the combination of IGFBP-2 and CEA could increase its sensitivity.
Stool biomarker	
Guaiac fecal occult blood test (gFOBT)	Use of the gFOBT as a CRC screening test reduced mortality by 11%-33% over 20 years of follow-up. However, it cannot distinguish upper GI bleeding from lower GI bleeding or human heme from non-human heme.
Fecal immunochemical test (FIT)	The FIT detects human globin using a human hemoglobin-specific immunoassay. It has higher sensitivity and specificity than gFOBT for detecting CRC and advanced adenomas.
Stool DNA (sDNA)	The Cologuard test (multi-target stool DNA test for CRC) showed higher sensitivity than gFOBT and FIT but also a higher false positive rate.
Stool miRNA	miR-92a, miR-21, miR144, miR-106a, miR17-92 cluster, miR135 were up-regulated in CRC, while miR-143 and miR-145 were down-regulated in CRC. However, none of the miRNAs showed adequate predictive value for use as a standalone CRC diagnostic test.

Table 3. CRC, colorectal cancer; TBT, Telomerase Biosensor Technology; mRNA, messenger RNA; HIF1A-AS1, hypoxia-inducible factor 1alpha-antisense RNA 1; AUC, area under the curve; CRNDE-h, colorectal neoplasia differentially expressed-h. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7206347/ [56]

2.1.5 Prognostic Biomarkers

The prognostic biomarker is defined as a biological molecule that provides information of patient's overall cancer outcome, independent of therapy [118].

1. Tissue Biomarkers

Recently, additional parameters have been used to specify the prognosis. Recent studies have highlighted the role of the DNA repair mechanisms in CRC and other malignancies and suggested potential role of associated factors as prognostic biomarkers [119].

A prominent example of DNA damage is represented by oxidative stress, where production of reactive oxygen species (ROS) leads to oxidative damage of DNA base

structure, formation of AP sites and breaks in DNA strands, all contributing to chromosomal instability. The ROS are prooxidative substances generated from diatomic oxygen. Their endogenous production occurs primarily in the mitochondria and they are involved within multiple cellular signalling pathways. The overproduction of ROS is associated with several pathological factors elevating risk of malignant transformation, such as chemical carcinogens, radiation, inflammatory bowel disease and obesity. Moreover, excessive ROS formation is observed in tumour microenvironment, including cancerous cells themselves. The accumulated damage due to high levels of ROS might lead to apoptotic pathways induction or necrosis in tumour cells. Thus, oxidative damage might be important in malignant transformation and early stages of tumour formation, as well as influence cellular proliferation and survival in the later stages of malignant process.

Under the physiological conditions, the oxidative damage is counterbalanced by activity of multiple pathways representing cellular DNA damage response. The various environmental factors associated with ROS overproduction interact with genetic variants within DNA repair pathways to modulate CRC susceptibility [120].

An important component of DNA repair mechanisms is base excision response pathway (BER), where activity of DNA glycosylases exerts removal of the bases marked by oxidative damage. The multiple DNA glycosylases active in human tissues include hOGG1, which drives excision of 8-oxo-dG, the highly mutagenic form of oxidized guanin. The activity of the hOGG1 is complemented with MUTYH, which is instrumental in removing the mis-incorporated adenines in opposite to 8-oxo-dG. Together, activity of hOGG1 and MUTYH prevents G:C > T:A transversion and enhances genomic stability.

We have recently reported decreased expression of hoGG1 and MUTYH genes in sporadic CRC tissues compared to healthy mucosa, suggesting decreased DNA repair capacity in tumor cells [121]. In another study [122], the BER-specific DNA repair capacity (BER-DRC) in non-cancerous mucosa was positively associated with OS and RFS of CRC patients. Furthermore, the prognosis was particularly favourable in subset of patients with high BER-DRC in non-cancerous mucosa with concurrent low BER-DRC detected in tumor tissue. In addition, the ratio of BER-DRC in mucosa and tumor tissue positively correlated with tumor staging. These results are consistent with the idea where decreased DNA repair capacity contributes to cancer pathogenesis, while,

on the other hand, the capacity of the DNA repair mechanisms also influences further survival and proliferation of the malignant cells, manifesting abundant ROS formation and oxidative damage. This is consistent with another study [123] showing that high BER genes expression in CRC tissue is associated with poor prognosis. Altogether, these results suggest that evaluation of DNA repair capacity might represent promising prognostic biomarker in colorectal cancer. Some additional parameters which have been used to determine the prognosis are mentioned below.

1) BRAF

The mitogen-activated protein kinase pathway, including RAS-RAF-MEK-ERK, is associated with migration, cell differentiation, angiogenesis, and proliferation. The dysregulation of this pathway leads to carcinogenesis. Approximately 8% of advanced CRC and 14% of localized stage II and III CRC cases have BRAF-activating mutations. KRAS and BRAF mutations are mutually exclusive. BRAF mutation significantly correlate with a tumor location in the proximal colon, poor differentiation, tumor size, and female sex. Patients with localized stage II and III CRC with BRAF mutations have also shown poorer OS rates compared to those with wild-type BRAF [30,124,125].

2) Microsatellite instability

Microsatellites are short and repetitive DNA sequences of 1–6 bp that can be found in both coding and noncoding regions. These regions are prone to replication errors that should be normally corrected by the mismatch repair (MMR) system [126,127]. Microsatellite instability (MSI) results from inactivation of the MMR genes through sporadic MLH1 promoter hypermethylation (80% of MSI CRC cases) or germline mutations in MMR genes such as MLH1, MSH2, MSH6, or PMS2 (20% of MSI CRC cases) [127,128]. The presence of deficient MMR leads to the aggregated somatic mutations and induces genomic instability, resulting in cancer-associated alterations. Lynch syndrome, also named hereditary non-polyposis colon cancer, results from germline mutations in MMR genes that lead to MSI. MSI with occurrence rate about 40% of right-sided CRC tumors is also associated with sporadic CRC. In sporadic MSI CRC, hypermethylation of the MLH1 promoter region causes MLH1 silencing. MMR status testing is recommended for patients with CRC for prognostic stratification. The MSI presence has been linked to better prognosis [128–130], which mostly reflects

robust immune response that has been shown to be triggered by MSI-related phenotype [131].

3) Cytosine preceding guanine island methylator phenotype

Cytosine preceding guanine (CpG) islands are regions that are common in promoter sites rich in CpG dinucleotides. Abnormal DNA methylation has been observed in almost all CRCs. CpG island methylator phenotype (CIMP) CRC (10%–20% of cases) have extremely high proportions of aberrantly methylated CpG loci [132,133]. Hypermethylation in CpG islands silences gene activity and results in dysregulated gene expression [112]. CIMP is often defined as hypermethylation of at least 3 loci in a selected panel of 5 genes (hMLH1, p16, MINT1, MINT3, and MINT31) that are associated with CpG islands [134,135]. The prognostic potential of this marker is yet to be fully elucidated, with existing studies mostly indicating relationship of CpG hypermethylation with poorer prognosis [136,137].

4) APC

Adenomatous polyposis coli (APC) gene mutation is associated with familial adenomatous polyposis and most sporadic CRC cases. APC has an important role in Wnt signaling, which is associated with cytoskeletal integrity, cellular proliferation, motility, and apoptosis by β-catenin regulation [138–140]. APC mutation can increase β-catenin levels, leading to increased c-myc expression, which is associated with cell proliferation. Therefore, APC mutation can cause the unregulated transcription of many oncogenes. It was recently reported that in advanced-stage CRC, patients with APC mutation and high miR-21 had shorter OS, indicating that APC mutation is associated with clinical outcomes of CRC [141].

5) p53

Fifty% to 70% of CRC have mutations in the TSG p53 [142,143]. When DNA is damaged, p53 plays a prominent role as a facilitator of DNA repair by halting the cell cycle to allow time for the repair machineries to restore genome stability. If DNA damage cannot be repaired, apoptosis is induced. Many studies have examined p53 mutations and their prognostic value in CRC patients. Some studies reported that p53

overexpression is linked with lower DFS, relapse-free survival (RFS), and OS rates [142,143,144].

6) SMAD4

SMAD4 is encoded by a gene located on chromosome 18. Alterations in the TSG SMAD4 can be categorized as loss of function of chromosome 18 gene function in protein expression. SMAD4 mediates the transforming growth factor-β superfamily signaling, which is frequently altered in human cancers. It is associated with cell proliferation, differentiation, apoptosis, and cell migration. A reported 30% to 40% of CRC cases show SMAD4 mutations. The loss of SMAD4 expression has been linked to worse clinical outcome [146,147].

2. Blood Biomarkers

1) CEA levels

Carcino-embryonic antigen (CEA) is a protein that is physiologically found on the cells of developing fetus and in pathological situation it is present on the surface of many types of cancer cells and high levels of CEA can be also detected in the blood of some cancer patients. Measuring CEA levels is recommended by the ASCO 2006 for the management of CRC patients. CEA testing every 3 months' post-surgery is recommended for patients with stage II or III CRC to detect possible disease recurrence. In addition to patient surveillance during the treatment, many studies have also recognized elevated preoperative CEA levels as a negative prognostic factor [148].

2) Neutrophil-to-lymphocyte ratio

The neutrophil-to-lymphocyte ratio (NLR) was first studied as a marker for immune responses to various stressful conditions. Lymphopenia is associated with impaired cell-mediated immunity, while neutrophilia is associated with systemic inflammation. Other studies found potential for NLR as a prognostic marker for pancreatic cancer, gastric cancer, and hepatocellular carcinoma. High level of the ratio (NLR >5) was significantly associated with elevated CEA level. Tsai et al. reported similar results after

analyzing 15 studies including 7,741 CRC patients. CRC patients with a NLR less than 5 before treatment were more likely to reach 5-year OS and DFS [149,150].

3) Concentration of cfDNA

A higher cfDNA concentration is reportedly related to significantly shorter OS in CRC patients. Furthermore, CRC patients with higher cfDNA levels showed a higher risk of disease recurrence. Similarly, Wang et al. observed APC, KRAS, and p53 mutations in the serum and found that patients with detectable cfDNA showed significantly higher rates of postoperative metastasis/recurrence than those without detectable cfDNA. Another study reported that patients with the KRAS mutation in the plasma and tissue showed shorter OS, indicating that the KRAS mutation in the plasma could be a prognostic marker for a poor outcome [151].

Studies have shown that the cfDNA concentration was reduced after primary resection, but upon CRC relapse, cfDNA levels dramatically increased. Another study found that when circulating tumor DNA was detected after CRC surgery, patients generally relapsed within 1 year [152]. These results indicate that postoperative cfDNA measurement can be used for monitoring the effects of therapy in cancer patients [153].

Summary of Prognostic Biomarkers in tissue and blood sample

Biomarker	Explanation
Tissue biomarker	
BRAF	Comprises the mitogen–activated protein kinase pathway, which is associated with cell differentiation, migration, angiogenesis, and proliferation. BRAF mutation is associated with poorer PFS and OS. The analysis of its prognostic role is recommended
MSI	MSI–high tumors have better prognosis than MSI–low tumors. MSI CRC patients showed longer OS and DFS than MSS CRC patients.
CIMP	The prognostic role of CIMP is unclear. However, the majority of studies reported that CIMP+/CIMP–high CRC patients showed poorer prognosis than CIMP–/CIMP–high CRC patients.
APC	APC is associated with FAP and most cases of sporadic CRC. APC mutation can cause unregulated transcription of many oncogenes Patients with APC mutation and high miR–21 reportedly had shorter OS.
p53	Some studies reported its association with lower DFS, RFS, and OS rates, but others reported no evidence of a prognostic role.
SMAD4	SMAD4 mediates the TGF–β superfamily signaling pathway, which is frequently altered in human cancers. Loss of SMAD4 was associated with poor DFS and OS.
Blood biomarker	
CEA	CEA level is reportedly significantly associated with patient outcomes. Preoperative CEA level was significantly associated with prognosis in patients with CRC metastasized to the liver.
NLR	Patients with elevated NLR were significantly associated with shorter OS and shorter PFS after treatment. CRC patients with a pretreatment NLR < 5 were more likely to have 5-year OS and DFS. Also, elevated pretreatment NLR was significantly related with poor OS and RFS in patients with liver metastasis.
cfDNA	CRC patients with higher cfDNA concentrations showed a higher risk of recurrence and shorter OS. Detectable APC, KRAS, and p53 mutations in the serum were significantly associated with a higher rate of postoperative metastasis/recurrence.

Table 4. PFS, progression-free survival; OS, overall survival; MSI, microsatellite instability; CRC, colorectal cancer; DFS, disease-free survival; MSS, microsatellite stable; CIMP, cytosine preceding guanine island methylator phenotype; FAP, familial adenomatous polyposis; RFS, relapse-free survival; TGF-β, transforming growth factor-β; NLR, neutrophil-to-lymphocyte ratio.
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7206347/ [56]

2.1.6 Predictive Biomarkers

The predictive biomarkers can be used to identify the response to a particular treatment. Until now, the most important indicator of response in patients with colorectal cancer corresponds to the DNA mismatch repair system deficiency [118]. The majority of chemotherapeutic agents exert their action to cause direct or indirect damage to DNA molecules, thereby substantially compromising survival and proliferation of cancerous cells [122]. The functional capacity of DNA repair pathways determines ability of tumour cells to overcome such damage and effective DNA repair mechanisms might contribute to chemoresistance.

Several chemotherapeutics, such as oxaliplatin, induce DNA cross-links, which may be associated with double-strand breaks (DSBs) formation. The MRN protein complex, consisting of MRE11, RAD50 and NBS1, is instrumental in initial stages of DSBs repair. Decreased expression of MRE11 and RAD50 in CRC cells was observed to predict better response to chemotherapy and longer progression-free survival [154].

Moreover, expression of MRE11 is regulated by activity of miRNAs, which was shown to modulate risk of CRC development [155]. In particular, miR-140 represents miRNA predicted to interact with MRE11 and with implied role in CRC development. In a recent study, we observed that the overexpression of miR-140 in CRC cells leads to decreased levels of MRE11 and increased accumulation of DSBs DNA damage marker γH2AX[156]. Furthermore, increased expression of miR-140 was associated with enhanced sensitivity of CRC cells to oxaliplatin. Notably, miR-140 did not influence oxaliplatin sensitivity in cell lines with suppressed expression of MRE11. These observations indicate that miR-140 downregulates MRE11, leading to decreased capacity for DNA DSBs repair and increased sensitivity to oxaliplatin genotoxicity.

The evaluation of DNA damage repair factors thus allows us to assess ability of cancer cells to counteract action of chemotherapy and might considerably contribute to treatment response prediction. Other important parameters that can be used as predictive markers are mentioned below.

1. Tissue Biomarkers

1) KRAS and NRAS

More than 50% of CRC cases show KRAS, NRAS, and BRAF mutations. The KRAS proto-oncogene encodes a GTPase protein (KRAS) that plays critical role in many molecular pathways, including the EGFR pathway. Only wild-type RAS tumors reportedly showed the clinical benefit of anti-EGFR antibody therapy such as cetuximab and panitumumab. KRAS mutations can be used as a negative predictive factor of a response to EGFR inhibitors [157,158].

2) Concomitant KRAS and BRAF mutations in colorectal cancer

BRAF mutation has been suggested as a potential predictive marker for anti-EGFR antibody therapy success. Although various studies reported poor prognosis in patients receiving anti-EGFR treatment and bearing BRAF mutation, there is ongoing investigation whether BRAF mutations could be used as true predictive marker. The results of a recent meta-analysis suggested that existing evidence remains inconclusive [159,160].

3) PIK3CA

PIK3CA mutation occurs in 10% to 20% of CRC patients [161]. PIK3CA encodes the p110α catalytic subunit of the class IA phosphatidylinositol 3-kinases (PI3Ks) and have important role in the RAS-mediated pathway that leads to proliferation, transformation, and tumor progression.

The results of a meta-analysis [162] indicate that patients with KRAS wild-type metastatic CRC and PIK3CA mutation have a lower response rate to anti-EGFR therapy and poor PFS and OS [163]. In particular, the PIK3CA mutation in exon 20 was significantly associated with a lack of the anti-EGFR treatment response. On the other hand, Karapetis et al. reported that PIK3CA mutation status cannot be used as a predictive marker for a benefit from anti-EGFR monoclonal antibodies. Their results showed that PIK3CA mutation was not associated with lower OS or PFS from cetuximab therapy in KRAS wild-type CRC patients [164].

Many studies have reported the protective effect of aspirin in CRC. A recent study reported that PIK3CA mutation could be a predictive marker for adjuvant aspirin

therapy in CRC patients. Liao et al. (2012) studied 964 patients with CRC and found out that PIK3CA-mutated CRC patients benefited from using aspirin in terms of longer cancer-specific survival and OS rates. In contrast, no such effect of aspirin treatment was observed in PIK3CA wild type CRC patients [165].

2. Blood Biomarkers

1) Cell-free DNA

Zitt et al. reported predictive effect of cfDNA concentration and integrity on efficiency of chemoradiotherapy for rectal cancer. The authors found that cfDNA concentration decreased in responders, whereas it increased in non-responders [166]. Agostini et al. reported that post-chemoradiotherapy levels of the cfDNA integrity index was significantly lower in rectal cancer patients who responded to the therapy. Similar results were observed in other cancer studies [167]. Patients who responded to the therapy showed decreased cfDNA, whereas those who did not respond to the therapy showed no change or increase in cfDNA [74,168].

Summary of Predictive Biomarkers

Biomarker	Explanation
Tissue biomarker	
KRAS, NRAS	KRAS proto-oncogene encodes a GTPase protein (KRAS) that plays an essential role in many molecular pathways including the EGFR pathway. Only wild-type RAS tumors showed the clinical benefit of anti-EGFR antibody therapy. KRAS mutations can be used as a negative predictive factor of a response to EGFR inhibitors. ,
BRAF	It has been reported that there is insufficient evidence that BRAF mutation could be used as a predictive marker for the benefit of anti-EGFR antibody therapy.
PIK3CA	Results are contradictory about whether PIK3CA mutation is associated with poor PFS and OS in KRAS wild-type CRC patient:
Blood biomarker	
cfDNA	The cfDNA concentration decreased after primary resection, but when the CRC relapsed, cfDNA levels dramatically increased. Similar results were found in rectal cancer patients who underwent chemoradiotherapy: The cfDNA concentration decreased in responders but increased in nonresponders.

Table 5. EGFR, epidermal growth factor receptor; PFS, progression-free survival; OS, overall survival; CRC, colorectal cancer; cfDNA, cell-free DNA. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7206347/ [56]

2.1.7 Diagnostics of CRC

New molecular CRC detection methods are currently being evaluated for detection of genomic mutations and alterations in protein expression that specifically correlate with the disease and that can help clinicians to make an accurate diagnosis. The global burden of colorectal cancer (CRC) is estimated to rise by 60%, with an incidence of

2.2 million new cases and 1.1 million deaths by 2030 [20]. The mortality can be largely attributed to metastatic cases, which account for approximately 90% of CRC related deaths. The high mortality rate largely results from the absence of an ideal method of screening [21]. Mainly since CRC is mostly asymptomatic until it progresses to advanced stages, the implementation of screening programs is required to reduce incidence and mortality rates with the aim at early detection of malignancy, or its precursor lesion in the early stage of cancer development, prior to the onset of symptoms, when treatment of cancer is the most effective [22,23]. The guidelines suggested that individuals aged 45 years and older with an average risk of CRC undergo regular screening [24].

The golden standard screening tests available is colonoscopy which have proven result in a decrease of CRC mortality by identification and removal of precancerous polyps. Despite of the high efficiency of colonoscopy, it is an invasive method, and some studies record the nonattendance rate of colonoscopies to be 10–20% after a positive fecal occult blood test, which decreases the total efficacy of the approach. The main factor related to non-adherence with colonoscopies are laxative bowel preparation and due to the lack of awareness of the significance of screening.

Fecal occult blood test (FOBT) and fecal immunochemical test (FIT) can be used as the best matrices for regular screening as these tests are more easily accepted by participants in population screening programs than tests like a colonoscopy, but positive results should be confirmed with timely colonoscopy. However, these methods are subject to various interfering factors with some causes of false-negative and false-positive results, also low sensitivity rates for detecting polyps [169,170]. In total, early, non-invasive, specific, and sensitive biomarkers are still required for screening strategies in colorectal cancer. Ideally, this aim can be achieved with a noninvasive and inexpensive method, using easily available biological samples. Moreover, despite the advances made over the last years, no single test is currently sufficient for CRC diagnosis.

The limitations of existing tests motivate extensive effort in development of novel, more efficient diagnostic tools. For example, high resolution Fourier transformation mass spectrometry have been done to report the overexpression of dipeptidase 1 (DPEP1) as a marker of high-grade CRC for screening of early neoplastic lesions and for prognostic stratification [17]. Blood-based biomarkers routinely present a considerable

benefit in the early diagnosis of CRC due to its availability and non-invasive collection with low cost and risk. However, biomarker detection in serum or plasma has some draw-backs such as costs for analyses, longevity of samples, laboratory errors and measurement errors [18]. Novel diagnostic methods such as analyzing of protein-based blood biomarkers are still under development. Several panels were suggested in individual studies. For example, Ivancic et al. analyzed blood from non-metastatic CRC patients by using targeted liquid chromatography and tandem mass spectrometry and proposed a panel of five proteins including leucine-rich alpha-2-glycoprotein 1, EGFR, inter-alpha-trypsin inhibitor heavy-chain family member 4, hemopexin, and superoxide dismutase 3 as novel colorectal cancer markers in detection of CRC [19,20].

Bhardwaj et al. conducted a study with serum samples and searched for potential biomarkers of adenomatous polyps and colorectal carcinomas through analyses of serum by a combination of high-performance liquid chromatography and mass spectrometry and performed five-marker blood-based profile consisting of mannan binding lectin serine protease 1 (MASP1), osteopontin, serum paraoxonase lactonase 3, transferrin receptor protein 1, and amphiregulin for the early detection of CRC [18,21]. Magnetic beads and matrix-assisted laser-desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry has been used by Yu et al. for the detection and validation of new biomarkers for the early detection of CRC [171]. In their study, 127 CRC serum samples and 90 healthy control serum samples have been analyzed. The protein serine/threonine kinase 4 (STK4 or MST1) was identified by tandem mass spectrometry (MS/MS) and validated with Western blotting and an enzyme-linked immunosorbent assay (ELISA) as the most altered between the groups. They demonstrated a downregulation of MST1 in CRC patients, with a sensitivity of 92.3% and specificity of 100% in the diagnosis of colorectal cancer when combined with carcinoembryonic antigen and FOBT. Their work also implied that MST1 could be a predictive marker for distant metastasis [171]. Despite the expansion of MS-based proteomics research and the large number of diagnostic biomarker candidates (Table 3 shows some examples of candidates for diagnostic protein biomarkers), none of them were yet successfully translated into clinical practice[56]. This probably occurs due to the difficulty of validating the possible biomarkers in large cohorts and comparing the results with the current screening methods. In addition, the methods are

demanding in terms of specialized equipment and trained personnel. However, the continuation of proteomic research is essential because, certainly, there is a space in the CRC screening that needs to be filled by reliable biomarkers. The increased understanding of the natural history of CRC and precursor lesions and the development and accumulation of evidence on screening technologies have supported the evolution of screening recommendations and implementation of CRC screening in clinical practice and public health programs [172,173].

2.1.8 Therapies for CRC

The aim of the CRC treatment is to achieve long-term remission whenever possible and increase the quality of life in the patients with advanced disease progression.

The CRC therapy requires multidisciplinary effort with specific therapeutic modalities including surgery, radiotherapy, chemotherapy and targeted biological therapy. Due to different etiopathogenetic and anatomical characteristics, treatment approaches differ between tumor locations in colon and rectum, respectively.

The ideal CRC treatment is to obtain complete removal of the tumor and metastases, which mostly requires surgical intervention. In the cases of very early cancer detection, the endoscopic approaches might be sufficient for complete removal of the malignant tissue.

Of note, in some cases, chemotherapy or radiotherapy might be applied before or after surgical resection, as neoadjuvant or adjuvant treatment, respectively, to maximally reduce and stabilize the tumor.

For the patients with unresectable tumor, the goal is a maximum shrinkage of the tumor and suppression of further tumor spread and growth, and radiotherapy and chemotherapy are the leading strategies for controlling disease in such patients. The palliative treatment also includes surgical procedures such as colostomy and tumor bypass surgery.

The chemotherapy in CRC treatment involves regimens where one or combination of the chemotherapeutic agents (cytostatics) are administered in particular number of cycles. The chemotherapeutic agents used include 5-fluorouracil, capecitabine, oxaliplatin and irinotecan, among others.

The targeted (biological) therapy is directed towards molecules and processes which are specifically associated with cancerous tissue. Therefore, a major advantage is sparing the healthy cells which is associated with considerably smaller extent of treatment side effects. The targeted therapy is combined with chemotherapy to improve the treatment outcomes. The various biomarkers can provide information helpful in evaluating the eligibility of patients for targeted therapy. In treatment of CRC, the targeted therapy includes inhibitors of angiogenesis (Bevacizumab, Regorafenib), and inhibitors of epidermal growth factor receptors (Cetuximab, Panitumumab). Immunotherapy is relatively new, emerging approach in systemic oncologic therapies.

The immunotherapy aims to boost immune system's response towards the cancerous cells and thus to counteract the cancer mechanisms of immune surveillance escape. The therapeutic modality includes monoclonal antibodies such as Pembrolizumab and Nivolumab. In CRC treatment, the immunotherapy seems to be particularly beneficial for patients displaying specific genetic tumor-associated characteristics, including mismatch-repair-deficient mutations and microsatellite instability [174].

2.2 Introduction to Alternative splicing

2.2.1 General introduction to AS

Alternative splicing of pre-mRNA (AS) is an important post-transcriptional regulatory mechanism that allows generation of distinct mRNAs from a single gene through the exclusion or inclusion of specific parts of pre-mRNA sequence (Figure 4). AS is highly regulated process associated with a vast array of normal and disease states and contributes to shaping tissue and cell identity via regulation of mRNA isoforms [175,176]. Emergent data demonstrated the various role of alternative splicing in cell differentiation, tissue specificity, developmental states, and disease conditions. Ninety two% to 94 % of human genes undergo AS and encode splice variants in the normal physiological processes, allowing the human genome to direct the synthesis of many more proteins than would be expected from its 20,000 protein-coding genes [177,178]. Alternative splicing pattern frequently changes under physiological conditions as a result of responding of organism to environment by determining which part of the genome and how it expresses, and AS regulation and deregulation is widespread in human diseases as well. In cancer associated phenotypes, AS is commonly changed

to avoid apoptosis, promote angiogenesis, or induce cell proliferation [179,180]. Moreover, up to one-fifth of human diseases has been estimated to be associated with aberrant splicing and an increasing number of splicing factor mutations is being discovered. The real number is likely even greater. Thus, as the AS is highly deregulated in various pathological states, the attempts are made to use alternative splicing as prognostic, diagnostic, or predictive biomarker. With the recent advancement in next-generation RNA sequencing (RNAseq), numerous potential biomarkers or therapeutic targets have been identified, including those involved in the progression or maintenance of multiple cancer types. Although currently many AS changes identified by transcriptome analysis lacks further experimental verification and are rarely used as biomarkers in CRC, alternative splicing is still considered as the current hot spot in cancer research, which continues to shed more light on various types of cancer and represents a promising source of new diagnostic and prognostic markers. Therefore, we need further research of AS to prove its potential as CRC biomarker [181,182].

In addition to its translational potential, exploring the regulation of alternative splicing and the underlying mechanisms adds important dimension of complexity to our understanding of the fundamental biology of both healthy and malignant cells.
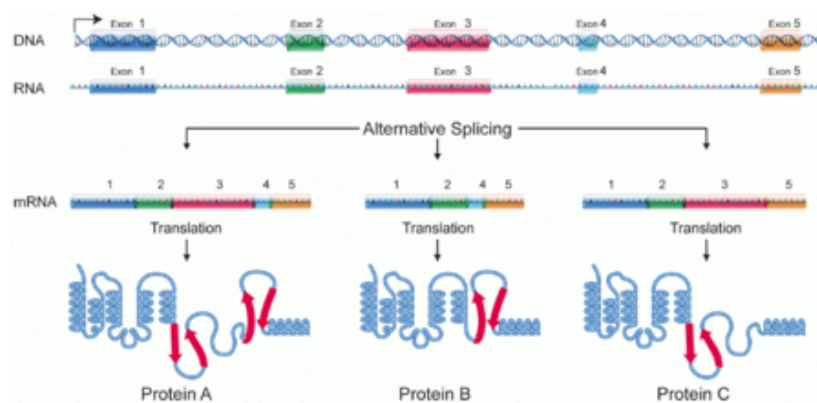

Fig 4. Schematic representation of alternative splicing.

Alternative splicing refers to the process by which a given gene is spliced into three protein isoforms. Copyright National Institutes of Health https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3933998/ [183]

2.2.2 Description of spliceosome and its function

The splicing of pre-mRNAs is known as vital process of gene expression in eukaryotes. Removal of introns from pre-mRNA sequences and ligation of exons are crucial steps in formation of protein-coding mRNAs and it is catalyzed through the activities of a large and dynamic RNA-protein (RNP) assembly called the spliceosome [184]. Higher intronic density and diversity in genetically complex organisms requires increased efficiency and accuracy of spliceosomes for pre-mRNA splicing. The machinery of spliceosome in humans, for example, consists of five uridine-rich small ribonucleoprotein (snRNP) complexes and of over than 200 proteins that are mostly conserved throughout the eukaryotic lineage [185,186].

For most series of splicing events, multiple splice sites are present on the nascent pre-mRNA. In this case, alternative splicing should not be viewed as a 'splice or not' case determined just by the innate properties of each splice site. Therefore, there is a competition where the splicing machinery must be differentiating and specifying between multiple splice sites in a context-dependent manner.

In human cells, splicing is essentially mediated by the spliceosome. The spliceosome consists of five small nuclear ribonucleoprotein complexes (snRNPs) – U1, U2, U4, U5, and U6, each containing a cognate U-rich noncoding small nuclear RNA (snRNA), and multiple protein factors, including NineTeen Complex (NTC). Spliceosome's active site is composed of RNA as identified from the crystal structure of a self-splicing group II intron [187]. After transcription, most snRNAs including U1, U2, U4, and U5 are transported into the cytoplasm for snRNP formation (Figure 5). A heptameric ring of Sm-proteins is convened onto each snRNA, which is followed by import into the nucleus, final maturation, and assembly of the spliceosome on the pre-mRNA [188]. The U6 snRNP is gathered into the nucleus and includes an Lsm ('like Sm') instead of Sm ring. In addition, less than 1% of human splice sites are processed by the much less abundant minor spliceosome, which possess the U5 snRNP, along with functional analogs of the remaining snRNPs in the major spliceosome (U11, U12, U4atac, and U6atac) [24].

Despite the small proportion of splice site substrates, the minor spliceosome has critical role in development [25,26]. Any dysregulation of the minor spliceosome is linked with multiple diseases, including neurodegeneration, developmental disorders [27-30], and cancer [32].

Moreover, the alterations within spliceosome lead to deregulated cellular splicing machinery, which might be associated with production of cancer-specific splicing isoforms. These isoforms and their encoded proteins may be participants of specific cancer hallmarks and contribute the effect on cancer cell phenotype. Aberrant alternative splicing should be added to the growing list of cancer hallmarks. Any inappropriate activity in the expression of splicing factors and splice factor kinases critically leads to change the alternative splicing of pre-mRNAs [189,190].
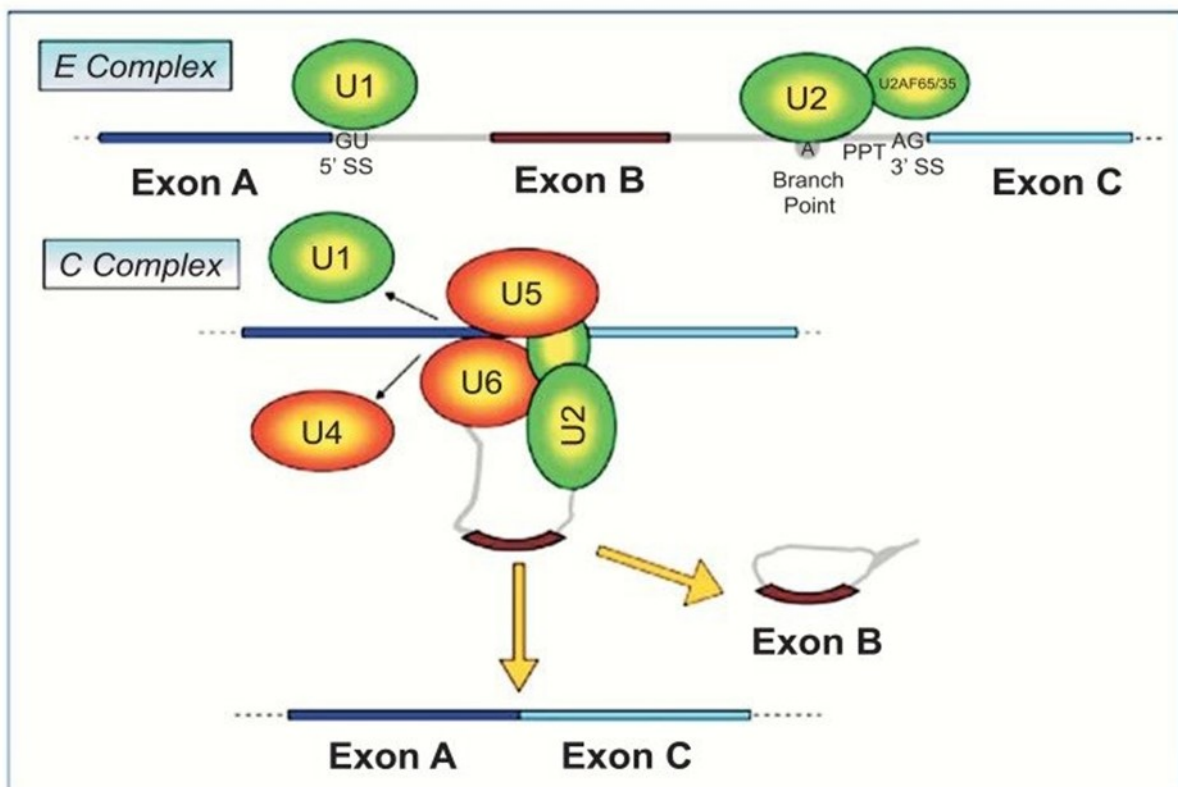
Fig 5. The alternative splicing mechanism. The spliceosome machinery (U1, U2, U4, U5 and U6) assembles on the nascent pre-mRNA. The conserved sequences that enable recognition of the mRNA by the spliceosome are: the 5' splice site (GU), the 3' splice site (AG), the branch point and the polypyrimidine tract (PPT). In E complex, U1 forms a base-pairing interaction with the 5'-splice site, whereas U2 base-pairs with the branch-point. Then, a tri-snRNP complex containing U4, U5 and U6 associates with the forming spliceosome, removing U1 and U4 (C complex). These steps allow the two transesterification reactions and join the exons.

https://www.researchgate.net/figure/The-alternative-splicing-mechanismThe-spliceosome-machinery-U1-U2-U4-U5-and-U6_fig1_263356754[191]

2.2.3   Types of alternative splicing

AS has been found in multiple tissues, stress conditions and developmental stages of eukaryotic organisms. In principle, standard pre-mRNA splicing is called constitutive splicing, and it is the particular splicing that leads to the most common mRNA variant of the given gene [192]. The majority of pre-mRNA splicing in eukaryotes belongs to this category (65%) [193]. In addition to constitutive splicing, alternative splicing can combine and ligate exons from a pre-mRNA differently, resulting in the production of various mature mRNAs with different sequence compositions. It has been estimated that approximately 92%-94% of human genes are alternatively spliced, implicating its significant role in the control of gene expression in higher eukaryotic cells [194].

Individual alternative splicing events are classified into several different subgroups: (i) exon skipping (ES) or cassette exons, where a whole exon is missed in comparison to the primary transcript; (ii) intron retention (IR), in which an intron can remain in the matured mRNA molecule; (iii) alternative acceptor (AA), where the 3′ splicing site is different, which is the result of the recognition of two or more splice sites at one end of an exon; and (iv) alternative donor (AD), the donor site, also known as 5′ splicing site [195,196]. Finally, based on a primary transcript there are three other additional AS complex events that give rise to alternative transcript variants which include (v) alternative splicing sites (ASS), where both donor and acceptor sites change; (vi) new intron (NI), when a splicing site appears in a reported exon; and (vii) retained exon (RE), a new exon replaces a previously annotated intron in the mature mRNA [197–199]. There are other, less frequent, complex events that give rise to alternative transcript variants, including mutually exclusive events, alternative transcription start sites, and multiple polyadenylation sites. Schematic representations of various types of AS events are described in Figure 6.
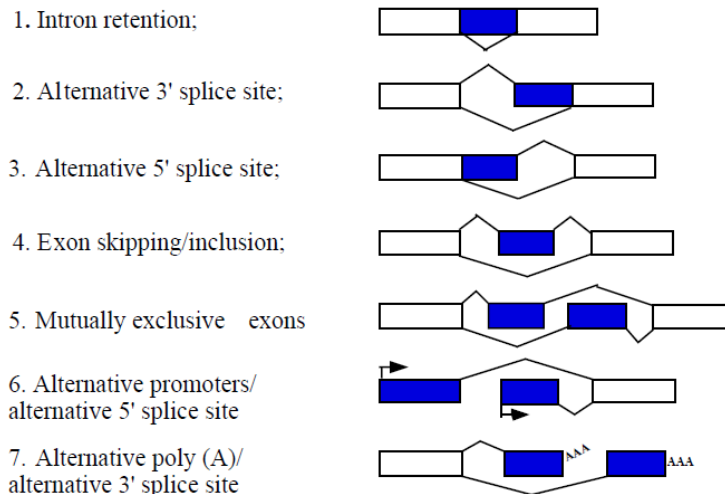
1. Intron retention;

2. Alternative 3' splice site;

3. Alternative 5' splice site;

4. Exon skipping/inclusion;

5. Mutually exclusive exons

6. Alternative promoters/ alternative 5' splice site

7. Alternative poly (A)/ alternative 3' splice site

Figure 6. Different patterns of alternative splicing. Modified based on http://www.jcb-jena.de/projects/alternative-splicing/

After discovery of alternative splicing, the regulation of gene expression was no longer confined only to on / off switches of transcription, but also extended to production of multiple variable mRNAs, with the capacity to produce alternative proteins, from a single transcription unit. Several mRNAs can be produced in a regulated, competitive manner, such that the reduced accumulation of one group of mRNAs can be compensated by the increased accumulation of another set of mRNAs, according to the demand of the cell type, or developmental stage of the organism. This finding added a new dimension to gene regulation.

2.2.4 Regulation of alternative splicing

Alternative splicing is the most clear-cut demonstration that represents a mechanism for gene regulation and expansion of the proteome, allowing for regulation of expression of specific isoforms in a developmental and tissue-specific manner. Regulation of alternative splicing is essential for a variety of biological processes including erythropoiesis, neuronal differentiation, and embryonic stem cell programming. The production of alternatively spliced mRNAs is regulated by a system of trans-acting proteins that bind to cis-acting sites on the primary transcript itself. The most widely studied factors regulating AS are proteins of the SR (Ser/Arg-rich) and hnRNP (heterogeneous ribonucleoprotein) families, in addition to several tissue-restricted AS factors [192,200]. These factors generally regulate AS by recognizing

cis-acting sequences in exons or introns and by promoting or suppressing the assembly of the spliceosome at adjacent splice sites. In contrast to these AS regulatory factors, much less is known about how or the extent to which components of the basal or "core" splicing machinery modulate splice site decisions. Most of the members of alternative splicing factor families have very high sequence identity to their family-members, have highly similar or identical RNA-recognition motifs, and recognize the same cis-elements. For example, Fox-1 and Fox-2 have very similar sequence, a 100% identical RRM domain and bind to the same RNA element and their function seems to be considerably redundant. The proteins regulating AS include splicing activators that promote the usage of a particular splice site, and splicing repressors that reduce the usage of a particular site. Mechanisms of alternative splicing are highly variable, and new examples are constantly being found, particularly through the use of high-throughput techniques [201,202]. It is of note that mechanisms of alternative splicing include RNA–protein interactions of splicing factors with regulatory sites termed silencers or enhancers, RNA–RNA base-pairing interactions, or chromatin-based effects that can change or determine splicing patterns [192,201]. The AS events are regulated in cell-type specific manner, so the expression of the individual splicing isoforms reflects their cell type- or conditions-specific function.

It appears likely that most splicing specificity is mainly specified by 3′ splice site (3′SS), 5′ splice site (5′SS) and branching point sequences, moreover by multiple cis-acting splicing regulatory elements (SRE) which promote splicing [203].

These splicing signals are classified in individual exons and are known as exonic splicing enhancers (ESEs) which increase the inclusion of exon into mature transcripts, while signals within exons that suppress splicing are studied as exonic splicing silencers (ESSs). Describing splicing enhancers and splicing silencers seems to be a big challenge because alternative splicing factors either enhance or suppress splicing based on the cis- elements position relative to the regulated exon. These signals have important role in the recognition of splice sites [204].

Over 300 splicing regulatory proteins are involved throughout the signaling process and following their binding to silencer sequences or splice enhancer, these proteins confer a positive or negative signal, influencing recognition of nearby splice sites by the spliceosome and therefore influencing splicing outcome. Deregulation of alternative splicing alters isoform ratios and lead to various cancers, muscular

dystrophies, and neurological diseases. The isoform ratios can be activated by differential expression of splicing factors which allows one pre-mRNA to code for multiple protein isoforms [205]. The importance of alternative splicing is illuminated by recent studies showing that how a particular alternative splicing factor regulates splicing on a genome wide scale. These alternative splicing networks create a very complicated mechanism of alternative splicing regulation. The combination of alternative splicing factors in one tissue at some point in development might be totally different from another tissue in another point in development, and the result of alternative splicing depends on the combination and ratio of alternative splicing factors interacting with one pre-mRNA. Additionally, alternative splicing factors can regulate either the alternative splicing or the mRNA stability and therefore protein abundance of other alternative splicing factors, thus creating a complex network of self-regulatory and cross-regulatory loops [192] Researchers hope to fully illustrate the regulatory systems involved in splicing. However, a better approach to understand the regulation of alternative splicing in a given condition requires a systems biology approach in which the expression status and targets of multiple if not all RBPs be assessed to start building these networks of co-regulated pathways, so that alternative splicing products from a given gene under particular conditions could be predicted by a splicing code [206,207]. Various families of splicing factors are able to bind to pre-mRNAs and modulate spliceosome recruitment, which lead to regulation of alternative splicing in cancer. Additionally, any mutations or changes in expression of splicing factors could lead to the activation of oncogenes and cancer pathways or the loss-of-function in tumor suppressors [188]. Splicing factors can be divided into hnRNPs, which typically suppress splicing, and Serine–Arginine (SR) proteins, that is known as important family of regulatory RBPs and function in both constitutive and alternative splicing [208]. The first SR proteins identified were SRSF1 (previously called SF2/ASF) and SRSF2 (previously called SC35). Interestingly, several key components of signaling pathways that relay extracellular signals to splicing factors that are typically deregulated in cancer can themselves be alternatively spliced to produce cancer-specific isoforms. For example, the SR protein family is often deregulated as the function of SR proteins tightly depends on their phosphorylation status, which itself is regulated by upstream kinases. For example, SRSF1 is itself phosphorylated upon activation of multiple signaling pathways, including the PI3K/AKT pathway while AKT signaling is primary activated in cancers, such as lung cancer. This result in constitutive phosphorylation

of SRSF1 and deregulated expression of Caspase 9a/9b. Another pathway involving an AKT-hnRNP U axis has also been shown to regulate Caspase 9a/9b ratio. This deregulated Caspase 9a/9b ratio has certain consequences on apoptosis and contributes to the ability of cancer cells to resist cell death. Diverse SR proteins tend to have a positive role in splicing regulation [209,210]. It is mentioned that SR proteins, for example SRSF2, are commonly mutated in a collection of neoplastic diseases or cancers of immature blood cells such as myelodysplastic syndromes (MDS). Mutations in SRSF2 that alter its sequence specificity on its target pre-mRNAs are more likely to be linked to MDS than nonsense mutations, indicating that a gain-of-function (binding to differential pre-mRNA targets) rather than loss-of-function of SRSF2 produces a new set of alternatively spliced mRNAs that are relevant to MDS development [211]. Of note, the alterations in splicing factors, whether due to mutations or altered expression, tend to have large effects on cell phenotype as these splicing factors bind to and regulate the splicing of hundreds of pre-mRNAs. Thus, cancer cells can alter the splicing of a large number of genes by deregulating a handful of splicing factors [212,213].

Altogether, spliceosomal regulatory factors that bind to RNA binding proteins (RBP) direct spliceosomes onto pre-mRNAs to perform alternative RNA splicing to generate more than one mRNAs from almost every gene in humans. Notably, the important function of intron diversity in regulation of gene expression and alternative splicing make the spliceosome an attractive new target for small-molecule, antisense, and genome-editing therapeutic interventions.

2.2.5 Alternative Splicing Regulation by MBNL

One of the important splicing regulators is MBNL family consisting of MBNL1, MBNL2, MBNL3 paralogs [205]. The study of MBNL proteins can aid our understanding of how concentrations of alternative splicing factors shift isoform ratios, and what properties define whether an isoform is sensitive to a various range of splicing factor concentrations. All members of the MBNL protein family are alternative splicing factors, which are subject to alternative splicing themselves. Thus, alternative splicing of primary transcripts from the three MBNL paralogous gives rise to nine, three and six protein isoforms of MBNL1, MBNL2 and MBNL3, respectively. The expression of MBNL proteins increases during development and they are considered to be negative

regulators of AS programs characteristic for embryonic stem cells [214]. Thus, their activity promotes differentiation and switch to adult-like AS patterns.

In general, MBNL1 and MBNL2 showed higher expression in brain, liver, kidney, pancreas, and muscle cell types. MBNL3 transcripts are expressed at much lower levels, with peak expression in placenta and no expression in skeletal muscles [215].

Both MBNL1 and MBNL2 have cytoplasmic functions in addition to being alternative splicing factors. In neural tissue, for example, cytoplasmatic MBNL promotes neurite morphogenesis [216]. However, it is not known how MBNL1 and MBNL2 separate their splicing functions when they are both abundant in the cell. Functional differences between MBNL1 and MBNL2 is that MBNL1 is the predominant splicing factor in muscle, while MBNL2 is the predominant splicing factor in brain and could have a general role in promoting proper expression of extra cellular matrix proteins and influences the localization of integrin alpha-3 to focal-adhesions in HeLa cells in cancer. MBNL proteins are expressed in both the nucleus and the cytoplasm, and the amount of protein in the nucleus versus the cytoplasm can undergo dynamical changes [204,217]. Therefore, the amount of total MBNL protein available for alternative splicing is just the nuclear fraction. Despite great advances in discovering the molecular functions of Muscleblind proteins, still so many questions remain unanswered. We know very little about the mechanism by which Muscleblind proteins regulate alternative splicing [218,219].

2.3   Alternative splicing in colorectal cancer

Colorectal cancer (CRC) is a severe risk to public health, and there is growing evidence that alternative splicing (AS) as the most extensively applied mechanism that account for the proteome diversity and cellular complexity plays a crucial role in cancer. AS is known as an important mechanism of regulating eukaryotic gene expression. Additionally, AS have a significant effect on the biologic characteristics of the final protein by adding or removing functional domains, controlling its location, modifying its protein-protein interactions. Splicing is a complex and accurate process, whose validity and integrity are defined by various factors, including the invariant 5'(donor) and 3'(acceptor) exon–intron junctions, exonic splicing enhancers, intronic splicing enhancers, exonic splicing silencers and intronic splicing silencers, different activities of trans-elements including SR proteins and heterogeneous nuclear

ribonucleoproteins (hnRNP). Errors in majority of components of alternative splicing processes were associated with diseases, including cancers, as proper regulation of AS have important role in both physiological and disease states of biological processes. In cancer research, many defects of AS or mutation and misregulation of splicing factors were linked to tumorigenesis, cancer metastasis and cancer drug resistance. Targeting AS or targeting splicing factors brings new therapeutic approaches in cancer diagnosis and prognosis markers to fight against cancer and may provide suitable biomarker candidates for CRC. Taken together, based on recent studies, in cancer, AS event can act as potential biomarkers for diagnosis, prognosis and potential new therapeutic targets in CRC [220,221].

2.3.1 Examples of Alternative splicing in cancer

In recent years the presence of AS in human diseases, especially in cancer, has been widely recognized. From the point of view of mechanism, aberrant AS events can directly affect major biogenesis and progression of tumors. For example, the systematic and coordinated alteration of AS was identified and in association with various precursor mRNAs could lead to alteration of the specific processes in carcinogenesis. Moreover, increasing evidence suggest AS as general and potent mechanism of transcriptional regulation [222,223]. Mutation of splicing regulatory elements on an exon or intron influence the splicing event in cancer. For example, in CRC, a point mutation of the APC tumor suppressor gene in cis-acting elements, such as ESE site, disturbs the binding sites of splicing factors such as SRp55, hnRNP A1 or ASF/SF2. In addition, mutations can produce new splicing regulatory elements [224–226]. Another example is a point mutation (G/A polymorphism) in intron 1 of the KLF6 gene, that results in formation of a new binding site for the SR protein SRp40. KLF6 act as a tumor suppressor gene and is somatically inactivated in prostate cancer. The new SRp40 binding site generates KLF6 splicing variants which antagonize the biological functions of wild type KLF6 [227]. There are many additional examples of alternative splicing isoforms with influence on human cancer including p53, fibronectin, fibroblast growth factor (FGF) receptor, vascular endothelial growth factor (VEGF) receptor, CD44, and murine double minute (MDM2) [228,229]. In addition, some of the alternatively spliced isoforms can assemble in tumors and are sufficient to promote cell

transformation and metastasis formation in cell culture. Alteration of SR factors and hnRNP proteins mostly occur in tumors and are conducted by alterations in the relative abundance of alternative splicing products, a typical signature of cancer cells. Cancer cell lines are characterized by a high level of alternative splicing events that are not conserved between human and mouse and are not found in physiological tissues [206].

Various aberrant splicing events have been linked to CRC pathogenesis [230]. For instance, inclusion of 1A exon during splicing of APC was observed to be 3.5 times more frequent in CRC cells than in non-affected tissue. This variant gives rise to premature stop codon, which leads to lower expression of this important tumor suppressor [220].

Another AS disturbance detected in CRC is represented by abnormal splicing patterns in KRAS. These include overabundance of antiapoptotic exon 4B containing isoform with respect to proapoptotic exon 4A containing variant [231].

Well described example of cancer-related AS is expression of various version of CD44. The alternative splicing of CD44, a trans-membrane glycoprotein involved in cell to cell and cell to matrix interactions, is a good example of how alternative splicing in cancer can be modulated by trans-acting factors. The production of different CD44 isoforms correlates with changes in the abundance of SR proteins. Various isoforms of CD44 with diverse inclusions in proximal extra-cellular domain of CD44 exons (v1-v10) have been identified. Wild type CD44 has no alternative exons and is predominantly expressed in normal tissues. However, CD44 isoforms, containing variant exons v5, v6 and v7, are over-expressed in various tumors and have been implicated in tumor cell invasion and metastasis [232].

Another example is represented by splicing factor SF2/ASF, which acts as a proto-oncogene and several target genes of SF2/ASF are essential for the oncogenic activity of S6K1, which is a novel oncogenic isoform of the mTOR substrate [233]. Furthermore, deregulation of RNA binding proteins in cancer is associated with alternative splicing of oncogenes or tumor-suppressor genes.

2.3.1   Alternative splicing and EMT

One of the major steps of cancer progression associated with extensive alternative splicing involvement is the process of epithelial-to-mesenchymal transition (EMT). The

EMT is highly conserved developmental program stimulated by the early steps of carcinoma metastases. During development, the EMT is known as an important process by which epithelial cells lose their epithelial morphology and acquire mesenchymal phenotype. EMT in adult tissues is controlled by the formation and particular structure of the extracellular matrix (ECM) components and ECM-remodeling matrix metalloproteinases (MMPs), as well as numerous growth factors or cytokines, including epidermal growth factor (EGF), hepatocyte growth factor (HGF), fibroblast growth factors (FGFs), and transforming growth factor (TGF)-β which is a potent activator of EMT [234,235]. EMT is involved in gastrulation via induction by members of the TGF-β family and results in the formation of the three multipotent germ layers [236]. In organogenesis, TGF-β is involved in EMT processes in the kidney, endocardial cushion, and subsequent atrioventricular valve formation. Effectively, all cell types are responsive to TGF-β, which is produced by various immune and nonimmune cells. TGF-β is also involved in proliferation regulation, migration, differentiation, and survival processes, and in cancer, depending on cellular context, can act either to suppress tumor growth or to activate tumor progression. Although TGF-β is key to normal inflammatory responses, its sustained expression during chronic inflammation stimulates fibrogenic processes and tumor promotion. Induction of EMT by TGF-β assumes a combination of Smad-dependent and Smad-independent events on cell junction complexes. The activation of Smad transcriptionally regulates key EMT associated factors, including Snail, Snail2 (Slug), Twist, and HMGA2. Smad-independent activity of TGF-β trigger phosphorylation of Par6 with a subsequent disorganization of the Par complex, leading to loss of cellular polarity. Multiple transcription factors have been identified in EMT, involved either in transcriptional inactivation of epithelial genes or activation of mesenchymal gene [237,238]. Among the transcription factors, Snail acts as central regulator of both developmental and pathological EMT, as Snail is critical for gastrulation and also has been associated with pathological conditions, such as fibrosis and cancer, in which its detrimental role is determined by its ability to induce EMT-like processes [239,240]. It directly suppresses E-cadherin, claudins, and other epithelial cell–cell adhesion molecules, and promotes expression of mesenchymal proteins, such as fibronectin and MMP-9. Increased expression of Snail mediates EMT program, induced for example by cytokines such as TGF-β, interleukin 6 and growth factors such as EGF or FGF. Expression of Snail is further regulated by an integrated and complex signaling network at the transcriptional

and posttranscriptional levels that includes integrin-linked kinase (ILK), phosphatidylinositol 3-kinase (PI3-K), mitogen-activated protein kinases (MAPKs), glycogen synthase kinase 3-β (GSK-3β), and NFκB pathways. The process of Epithelial-Mesenchymal Transition (EMT) contributes towards the exacerbation of cancer invasiveness along with the obtaining of stem like and the subsequent drug resistant behavior [241,242].

In addition to usual transcriptional regulation of EMT that has been largely studied at the level of altered gene transcription or posttranslational modifications, it is becoming increasingly clear that alternative splicing processes provide an additional layer of gene regulation that is critical in shaping the EMT process, particularly in cancer progression. The dysregulated AS might represent important component of cancer-related EMT induction and thus critically influence disease progression. Notably, the splice variants induced by aberrant AS of the critical genes might promote the metastatic potential and stem-like/chemo resistant behavior of cancer cells, further facilitating tumorigenesis.

Numerous splicing alterations implied in the process are controlled by various splicing factors including ESRP, RBFOX2 and MBNL proteins.

Epithelial splicing regulatory proteins 1 and 2 (ESRP1 and ESRP2), which act as splicing regulators of genes associated with epithelial cell phenotype, have been implicated as central coordinators of AS networks in EMT. For example, AS of CD44 with a switch from a variant isoform (CD44v) to a standard isoform (CD44s) is promoted by downregulated ESRP1 and is strongly related to the EMT process [243]. CD44s expression is up-regulated in breast tumors and correlates with levels of the mesenchymal marker N-cadherin and breast cancer progression. Down-regulation of ESRP1 expression, using small hairpin RNA (shRNA), led to a switch from CD44v to CD44s. Conversely, overexpression of ESRP1 increases CD44v expression and prevents cells from undergoing EMT.

Another EMT-associated AS event arising from ESPR1 downregulation is exon 4 skipping during splicing of TCF4, an important component of Wnt signaling pathway. This subsequently leads to enhanced Wnt signaling related to EMT mechanisms [244].

Taken together, EMT represents both an essential step in development and normal morphogenic processes, as well as a mechanism that facilitates pathological processes such as fibrosis and chronic inflammation and many stages of tumor

progression. Accordingly, differentiating between how the EMT-related AS patterns differ between normal development and disease has become an important area of investigation [245].

3   Goals and hypotheses

In this study we analyzed gene expression of the MBNL family of regulators of alternative splicing, together with the MBNL-target splicing events in FOXP1 and EPB41L3 genes and tumor-related CD44 variants. We chose a patient cohort in various stages of colorectal cancer development. After RNA isolation and reverse transcription, the relative gene expression of a selected gene panel was tested by quantitative real-time PCR, followed by statistical analysis. Our goal was to identify deregulation of MBNL genes in colorectal cancer and correlate the expression levels with clinic-pathological characteristics. We conjectured that down-regulation of their transcripts in tumor tissue compared to matched non-tumor mucosa can lead to transition of alternative splicing patterns towards a less differentiated phenotype. This would further highlight the importance of alternative splicing regulation for tumor growth and propagation. Moreover, we hypothesized that the lower expression of MBNLs would be associated with worse disease progression.

We aimed for comparison of deregulation of MBNL expression in tumor tissue with non-tumor mucosa. In addition, we aimed at analyzing of expression changes in selected genes panel including FOXP1, EPB41L3, CD44 and its cancer-related variants 3 and 6.

In collaboration with Institute of Experimental Medicine we addressed whether main glycosylases involved in base excision repair (BER) and in removing oxidative DNA damage may exert the potential to become biomarkers of CRC risk and progression. Furthermore, we investigated on CRC cells a role of miRNA-140 in regulating the initial step of Homologous recombination DNA repair (HR), mediated by *MRE11E*, responsible for triggering the removal of double-strand breaks and inter-strand crosslinks, induced by chemotherapeutic oxaliplatin. These may represent first step in dissecting biomarkers of treatment prediction.

3.1 Hypothesis 1: All three MNBL paralogs in CRC are deregulated by expression of their mRNA in CRC tissue

The MBNL proteins are essential regulators of alternative splicing and aberrant splicing is implied in cancer pathogenesis. Accordingly, deregulated MBNL activity has been linked to various malignancies, including colorectal cancer. However, no study so far has investigated activity of the all the MBNL paralogs in colorectal cancer.

We aimed to examine possible deregulation of all three MNBL paralogs in CRC by comparing expression of their mRNA in CRC tissue with expression levels detected in control sample.

Conclusion: Indeed, we detected significantly downregulated expression of all three MBNL paralogs in CRC tissue.

3.2 Hypothesis 2: CRC is associated with deregulated expression of specific splice variants of selected genes (FOXP1, CD44, EPB41L3)

A potential deregulation of MBNL proteins would indicate abnormal splicing patterns in malignant cells. We hypothesized that this would lead to deregulated expression of specific splice variants of genes undergoing alternative splicing.

We thus additionally analyzed CRC-associated expression of alternatively spliced variants of target genes shown to be regulated by MBNL or genes displaying abundant alternative splicing (EPB41L3, FOXP1 and CD44 isoforms). We hypothesized that expression of specific isoforms would be congruent with any potential change in MBNL expression in CRC tissue.

Conclusion: Yes, we detected significant deregulation of selected splice variants.

3.3 Hypothesis 3: Deregulated expression of the studied genes is related to clinical parameters (T, N, M, G, UICC, overall/recurrence-free survival)

We further aimed to study the expression of investigated genes in context of clinical characteristics and to investigate their potential as prognostic and predictive biomarkers. Considering the relevance of the alternative splicing deregulation for cancer pathogenesis, we hypothesized that extent of alternative splicing disruption

might be related to disease progression. We analyzed expression of MBNL genes and the genes in target panel in association with clinical parameters (T, N, M, G, UICC), as well as overall/recurrence-free survival.

Conclusion: We did not detect any association between expression levels of MBNL paralogs and clinical parameters in CRC group. We found several significant associations between expression of specific variants in target gene panel and clinical characteristics.

3.4 Hypothesis 4: Colorectal cancer is associated with deregulation of base excision repair glycosylases hOGG1 and MUTYH

The base excision repair glycosylases hoGG1 and MUTYH are important components of DNA damage response pathways. Unrepaired oxidative damage and resulting mutations in tumor suppressor genes and proto-oncogenes might contribute to cancerogenesis. We thus investigated expression of hOGG1 and MUTYH in colorectal cancer tissue and adjacent mucosa. We hypothesized that colorectal cancer is associated with down-regulated activity of hoOGG1 and MUTYH.

Conclusion: We detected decreased expression of hOGG1 and MUTYH in colorectal cancer cells compared to adjacent non-malignant mucosa.

3.5 Hypothesis 5: Expression of miRNA-140 is deregulated in colorectal cancer cells and modulates sensitivity to oxaliplatin treatment

We next studied miR-mediated regulation of MRE11, a factor involved in double-strand break repair, in CRC. *In silico* analysis indicated miR-140 as a molecule targeting MRE11 with strong association to progression-free survival. We hypothesized that CRC is associated with deregulation of miR-140. Further, since miRNA tend to suppress expression of their target genes, we hypothesized that increasing miR-140 expression *in vitro* will lead to suppressed expression of MRE11. Moreover, miRNA-

regulated activity level of DSB repair system involving MRE11 is expected to modulate sensitivity to oxaliplatin.

Conclusion: We detected downregulated miR-140 in CRC tissues. Increasing miR-140 in CRC lines *in vitro* resulted to suppressed expression of MRE11 and increased sensitivity to oxaliplatin treatment.

4   Materials and methods

In the present study, we analyzed the expression of three MBNL paralogs in colorectal cancer and selected variants of forkhead family transcription factor (FOXP1), Erythrocyte Membrane Protein Band 4.1 Like 3 (EPB41L3) and CD44 transcript variants 3 and 6. To test the effect of MBNL activity on the alternative splicing of the target genes, we selected FOXP1 and EPB41L3 exons described to be regulated by MBNL1 as reported by Hong Han et al [246]. CD44 transcript variant 3 and CD44 transcript variant 6 were selected as particular isoforms of CD44 related to cancer diseases. Additionally, to serve as potential marker for MBNL functional status, analysis of FOXP1, EPB41L3 and CD44 alternative splicing could bear important information given their own oncogenic potential. Selection of reference genes – hypoxanthine phosphoribosyltransferase1 (HPRT1), Histidyl-TRNA Synthetase (HARS) and actin beta (ACTB) was based on existing literature data and our own experience.

Primer design

Quantitative real-time polymerase chain reaction (PCR) primers for the selected genes were designed using Primer3 software with special emphasis on exon/exon-junction-spanning assays. In case of special splicing variants, primers were designed to align into the exon of interest, common assays were designed on exons typical for all other variants included in NCBI RefSeq database at the time of primer design (fall 2016). Designed primers were aligned to non-redundant mRNA database using BLAST to check putative amplification of off target sequences and to validate priming of desired variants. Primer sequences used in the study are summarized in Table 6.

List of used primers and their characteristics

Gene name	Primer type	Primer sequence [5' - 3']	Amplicon size [nt]	GC content [%]	Predicted melting temperature [°C]
ACTB	forward	CCAACCGCGAGAAGATGA	97	56	60
	reverse	CCAGAGGCGTACAGGGATAG		60	59
HPRT1	forward	TGACCTTGATTTATTTTGCATACC	102	33	59
	reverse	CGAGCAAGACGTTCAGTCCT		55	60
HARS	forward	ATTGGGGTGGAGCGGATTTT	78	0,5	60
	reverse	TGTCTCCGTGGTCCGTATCT		0,55	60
MBNL1	forward	ATCGCCTGCTTTGATTCATT	76	40	59
	reverse	TTAAATGTGGGGGTGGATGA		45	60
MBNL2	forward	GCCCAGCAGATGCAATTTAT	137	45	60
	reverse	AACCCAACTCCAGGGGTTAC		55	60
MBNL3	forward	ATTCCTGGAAACCCACCTCT	89	50	59
	reverse	TGAAATTCTCGGCAAACCTC		45	60
FOXP1com	forward	CACGTGGAAGAATGCAGTGC	88	55	60
	reverse	CACTGTCCATACTGCCCCTT		55	59
FOXP1var	forward	ACGTGGAAGGGTGCCATTC	114	58	60
	reverse	GCGGCCACGTTTAAACTCTT		50	59

CD44	forward	CCCAGATGGAGAAAGCTCTG	113	55	60
	reverse	GTTGTTTGCTGCACAGATGG		50	60
CD44v3	forward	TCCCTGCTACCAATATGGACTC	122	50	60
	reverse	ACTCTGCTGCGTTGTCATTG		50	60
CD44v6	forward	AGGAACAGTGGTTTGGCAAC	68	50	60
	reverse	CGAATGGGAGTCTTCTTTGG		50	60
EPB41L3com	forward	AAAGAGGCCAAAGAGCAGCA	115	50	60
	reverse	GCAAGCTAAGTTATTCCTCTGGTC		46	59
EPB41L3var	forward	ACCATGACCAGGAATAACTTAGCT	101	42	60
	reverse	AGTCAGTTGGGTTAGAAGAGGG		50	59

Table 6: List of used primers and their characteristics. Primers designated as var. are specific for particular splice variant know to be regulated by MBNL function, primers labeled as com are amplifying other mRNA splice isoforms except the specific MBNL regulated variant.

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8126336/[246]

4.1   Sample collection and processing

In this retrospective study, 108 patients from the Czech Republic with clinically confirmed CRC at the Department of Surgery of University Hospital in Pilsen between years 2008 to 2016 were analyzed. Clinical data were collected since 2008 to 2017 by clinicians and anonymized before transfer for statistical analysis.

During each patient's surgery, tissue was collected within 20 min after the removal of the tumor tissue from the patient and non-tumor mucosa (anatomically the most distant tissue that was macroscopically intact, in the range of 15–20 cm from the tumor location). Tissue samples with maximal size of 5 mm in the longest axis were gathered,

immediately frozen in cryotubes (Thermo Fisher Scientific, Waltham, MA, USA) and stored at -80˚C.

All patients and donors enrolled in the study agreed to the processing of their blood samples by signing informed consent. The study protocol was approved by the Ethical committee of Faculty of Medicine and University Hospital in Pilsen and complies with the International Ethical Guidelines for Biomedical Research Involving Human Subjects. All blood samples were collected by clinical personnel of the University Hospital in Pilsen and patient data has been anonymized. An overview of cohort characteristics is presented in Table 7.

Characteristics of the analyzed colorectal cancer patients.

Characteristic	Category	Number	%
gender	F	44	40.7
	M	64	59.3
age (in years)	30–50	7	6.5
	50–70	57	52.8
	>70	44	40.7
T stage	T1	3	2.8
	T2	23	21.3
	T3	66	61.1
	T4	14	13.0
	unknown	2	1.9
N stage	N0	62	57.4
	N1	21	19.4
	N2	21	19.4
	N3	1	0.9
	unknown	3	2.8
M stage	M0	79	73.1
	M1	26	24.1
	unknown	3	2.8
tumor grade	G1	13	12.0
	G2	74	68.5

	G3	14	13.0
	unknown	7	6.5
UICC staging	I	23	21.3
	II	33	30.6
	III	27	25.0
	IV	24	22.2
	unknown	1	0.9
tumor location	right or transversum	41	38.0
	left or sigma	30	27.8
	rectum or rectosigma	36	33.3
	non-specific	1	0.9

Table 7. F: Female; M: male; tumor classification based on TNM classification on malignant tumors with following categories: T stage: size of the primary tumor; N: degree of tumor spread in lymph nodes; M: presence of distant metastasis; G: histological tumor grade; UICC staging: tumor classification according to Union for International Cancer Control. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8126336/[246]

4.2   Isolation of RNA and quality validation

Standard protocol for RNA isolation using TriReagent RT (MRC, USA) was used to isolate RNA from frozen samples.

1.Isolate RNA

a. The tissue was removed from the cryotube and placed in liquid nitrogen in a mortar.

b. The tissue was pulverized and transferred into an Eppendorf tube with 1ml TriReagent RT chilled to 4 ˚C

c. Add 0.5 mL of isopropanol to the aqueous phase, per 1 mL of TRIzol™ Reagent used for lysis.

d. Incubate for 10 minutes at 4°C.

e. Centrifuge for 10 minutes at 12,000 × g at 4°C. (Total RNA precipitate forms a white gel-like pellet at the bottom of the tube).

 f. Discard the supernatant with a micropipettor.

2.Wash the RNA

a. Resuspend the pellet in 1 mL of 75% ethanol per 1 mL of TRIzol™ Reagent used for lysis.

 b. Vortex the sample briefly, then centrifuge for 5 minutes at 7500 × g at 4°C.

 c. Discard the supernatant with a micropipettor.

 d. vacuum or air dry the RNA pellet for 5–10 minutes.

 3.Solubilize the RNA

 a. Resuspend the pellet in 20–50 µl of nuclease-free water (Ambion, Carslbad, CA, USA).

 b. Incubate in a water bath or heat block set at 55–60°C for 10–15 minutes.

 4. RNA concentration was assessed by absorbance measurement using the Infinite M200 (Tecan, Männedorf, Switzerland) in the NanoQuant setting.

5. RNA quality was assessed by agarose gel electrophoresis. Only samples with a 230 nm/260 nm ratio >1.7 and samples with proper bands without degradation on agarose gel electrophoresis were used for further analysis.

4.3   Reverse transcription (DNase, PCR control)

cDNA was synthetized according to RevertAid First Strand cDNA Synthesis Kit manufacturer's protocol (THERMO SCIENTIFIC, Waltham, MA, USA).

4.3.1 Preparation

1.Place RevertAid First Strand cDNA Synthesis Kit components and RNA samples on ice.

2. Mix and then centrifuge briefly to collect contents at the bottom of the tube.

4.3.2   Reaction

1. Determine the number of reactions required, including controls. Calculate the volumes of each component required for all reactions (allow 10% extra for pipetting errors)

2. Using 500 ng of total RNA in 20 µl reaction for priming of reverse transcription reaction, a combination of oligo(dT)18 and random hexamer primers, each at 2.5 µM final concentration, was used.

3. Primer mix and RNA mixture is inserted into the cycle for 5 min in 75°C Then the reaction mixture is added and put in program in BioRad T100 for 60 min / 42 °C and 5 min / 70°C and ∞ min / 4°C. After completion of the test tube stored at -80 ° C.

4. The quality of cDNA and possible contamination by genomic DNA was assessed by control PCR reaction (GAPDH amplification, 40 cycles) and agarose gel electrophoresis.

4.4   Quantitative real-time PCR

The Quantitative real-time PCR is known as a reliable screening method for biosignatures in clinical specimens with a high sensitivity for detection.

Quantitative real-time PCR with 5x HOT FIREPol® EvaGreen® qPCR SuperMix (Solis BioDyne, Tartu, Estonia) was used for the quantitative PCR. cDNA was diluted 100x and 2 µl of cDNA were used in each reaction, having a final volume of 10 µl. Optimal cycling parameters and annealing temperatures were assessed by the measurement of sensitivity, specificity and efficiency of individual quantitative PCR reactions. CFX 96 quantitative real-time PCR cycler (Biorad, CA, USA) was used for the analysis.

Cycling parameters were: initial hold at 50°C for 20 s and initial denaturation at 95°C for 12 min followed by 45 cycles consisting of denaturation at 95°C for 15 s and annealing and extension at 60°C for 60 s. Results were analyzed and basic statistical analysis was carried out using the Bio-Rad CFX Manager software.

4.5   Statistical analysis of data

Standard frequency tables and descriptive statistics were used to characterize the patient group. Relative expression levels of measured genes in the tumor tissue

(relative to adjacent non- tumor tissue) were expressed by $-\Delta\Delta C_t$ values and analyzed using parametric tests after ensuring that the data can be modeled using normal distribution (histogram plots and Shapiro-Wilk test). We employed one sample t-test against zero reference to test for significant up- or downregulation of the respective expression of genes. Associations between the expression levels and clinical parameters (T, N, M, G, UICC) were estimated using two-sample t-test (difference between two groups), one-way ANOVA (testing for non-uniform means in multiple groups) and Kendall's tau (testing for increasing/decreasing trends across multiple ordinal categories).

Recurrence-free survival (RFS) was estimated for time period from the date of surgery until the date of first documented disease recurrence or death. Overall survival (OS) was determined from the date of anti-EGFR therapy initiation until the date of death, regardless of its cause. Patients who had not progressed or died were censored at the date of last follow-up in 2020. Significance of associations between gene expression and survival times was assessed using univariable Cox proportional hazards model. In order to visualize these associations and detect possible non-proportional ones, significant or borderline results were reviewed using automated stratification. In this procedure, a threshold (cut-off) expression $(-\Delta\Delta C_t)$ value was determined for each gene by an automated optimization process finding the threshold providing the lowest Cox-Mantel *p* value in two-sample Kaplan-Meier survival analysis, which was finally verified using Gehan-Wilcoxon test.

We report two-tailed *p*-values with respective statistical significance level set at $\alpha = 0.05$. To control for false discovery rate (FDR), we applied the Benjamini-Hochberg procedure over all the significance test performed in the study. The estimated FDR is 10%, which corresponds to 90% of the presented significant results to being true positives. A conservative overall FDR of 5% would be achieved after adjusting significance level to 0.022. Statistical analysis was performed using STATISTICA software system (StatSoft, Inc.2013, Version 12,www.statsoft.com) and Matlab (2019b, MathWorks Inc., Natick, MA, USA).

4.6 Study 2 Methods

The study involved tumor tissue and adjacent mucosa samples from 193 (122 males, 71 females; median age at diagnosis 69.5 years) sporadic CRC patients.

DNA was extracted from tumor and adjacent non-malignant mucosa, using the DNeasy Blood and Tissue Kit (Qiagen, Courtaboeuf, France), followed by amplification of coding sequences as well as adajacent intron regions of *MUTYH* gene. The resulting amplicons were prescreened by high resolution melting analysis. The samples with non-standard melting curves were subsequently sequenced by a modified Sanger method.

The mutations in KRAS gene were detected using SNaPshot analysis, targeting mutations in codon 12 and 13 at nucleotides c.34, 35, 37 and 38, corresponding to approximately 97% of all known variants.

To evaluate expression of hOGG1 and MUTYH, RNA was isolated using TRIzol (Invitrogen, Waltham, MA, USA) and cDNA was synthesized from 0.5 µg of RNA using the RevertAld™ First-strand cDNA synthesis kit (Fermentas, Burlington, ON, Canada). The preamplified cDNA was sequenced by qPCR (BioMark™ HD System; Fluidigm, San Francisco, CA, USA). TOP1 was used as a reference gene.

For imunohistochemical analysis of hOGG1 protein, the samples were fixed with 10 % formalin solution and processed into three to five µm-thick slices. The first slices were stained with haematoxylin-eosin and the second slices with 8-hydroxyguanine DNA glycosylase (OGG1) polyclonal antibody (PA-116505, Thermo Fisher Scientific, Waltham, MA, USA). Subsequently, the slices were visualized using the LSAB+System HRP kit (streptavidin-biotin peroxidase detection kit, DAKO, Glostrup, Denmark).

4.7 Study 3 Methods

Paired tumor and non-malignant adjacent mucosa samples from 50 CR patients were analyzed.

Bioinformatics analysis included data from TargetScan and miRNA-Seq transcriptional profiles accompanied with clinical information obtained from TCGA

(https://portal.gdc.cancer.gov; projects TCGA-READ (rectal adenocarcinoma, n = 155) and TCGA-COAD (colon adenocarcinoma, n = 476)).

Human colorectal cancer cell lines HCT116, DLD1, and HT29 were cultured in Dulbecco's modified Eagle's medium (DMEM) with 10% fetal bovine serum, 1 mM of l-glutamine 1 mM of sodium pyruvate and 1 mM of penicillin/streptomycin. All cells were cultured in a humidified incubator at 37°C, with 5% CO2.

Cells were transfected in 6-well plates at 60%–80% confluency with 2.5 pmol of MISSION miRNA hsa-miR-140-3p miRNA Mimics or with Negative Control miRNA Mimics with no homology to the human genome using Lipofectamine® RNAiMAX 2000 (Invitrogen™). All the experiments in cell lines were performed in three independent repeats. The efficiency of transfection was analyzed by qPCR measuring expression levels of transfected miRNAs as compared to negative controls.

Forty-eight hours after transfection, total RNA was extracted. The concentration of the total RNA was measured by Nanodrop™ 8000 Spectrophotometer and the integrity of mRNA (RNA integrity number (RIN)) of each sample was determined by Agilent RNA 6000 Nano Kit by Agilent Bioanalyzer 2100. Reverse transcription was performed using the High-Capacity cDNA Reverse Transcription Kit.

Expression levels of miR-140 were measured using TaqMan MicroRNA Assays at 7500 Real Time PCR System . The reaction contained 2 μl of a sample with 40 ng of cDNA, 10 μl of TaqMan™ Universal PCR Master Mix, 1 μl of the assay, and 7 μl of RNAse-free water. The thermal protocol was as follows: 50°C for 2 min, 95°C for 10 min, 40 cycles of 95°C for 15 s, and 60°C for 60 s plus melting curve analysis. MiRNA expression was normalized to RNU6B, and all data were subsequently analyzed by the $2^{-\Delta\Delta Ct}$ method.

Oxaliplatin was dissolved in dimethyl sulfoxide at the concentration of 100 mM and stored at 4°C. To assess the chemosensitivity of CRC cells with overexpressed miR-140 and control cells, both cells were treated with a 6 μM concentration of oxaliplatin 24 h after miRNA mimics transfection and analyzed for cell viability.

For clonogenicity formation assay (CFA), 48 h after cell transfection with miRNA mimics, 500 cells per well were plated for colony formation assay onto 6-well plates and cultured in DMEM. Twelve days later, colonies were fixed with 3% formaldehyde, stained with 1% crystal violet, and counted.

For proliferation assay, cells were plated onto 96-well plates at a density of 3 × 104 cells per well. The metabolic activity of the cells was measured 24 h after plating by adding WST-1 solution into the media. Absorbance at 450 and 690 nm was measured on BioTek ELx808 absorbance microplate reader.

Cells were seeded on 12 well plates, harvested, washed with PBS, and centrifuged at 1,000 rpm for 10 min. Then, 1 ml of propidium iodide (PI) staining solution was added to the cell pellet, and cells were incubated for 30 min at 37°C in the dark. After incubation, samples were analyzed using a flow cytometer. Measured data were evaluated with FlowLogic software (Inivai Technologies, Mentone, VIC, Australia).

Proteins (20 µg) were loaded and separated in 10% sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) gels at 15 mA for 60 min. Then, the separated proteins were transferred to 0.45 µm Amersham Protran Nitrocellulose Blotting Membrane in methanol transfer buffer using Mini Trans-Blot Cell. The membranes were blocked with 5% bovine serum albumin (BSA) in Tris-buffered saline containing Tween 20 (TBST; 20 mM of Tris–HCl, pH 7.4, 0.15 M of NaCl, and 0.1% Tween 20) for 1 h and incubated with anti-MRE11, anti-γH2AX, anti-RAD51 (Cell Signaling, Leiden, the Netherlands) and anti-GAPDH antibodies (Abcam, Cambridge, UK) at 4°C overnight, followed by incubation with goat anti-rabbit secondary antibody conjugated with horseradish peroxidase (Abcam, Cambridge, UK). The membranes were then incubated with Immobilon Western Chemiluminescent HRP Substrate (EMD Millipore Corporation, Billerica, MA, USA) and visualized by Azure c600 (Azure Biosystems, Dublin, CA, USA).

For the preparation of recombinant lentiviruses expressing MRE11 shRNAs, HEK293FT cells seeded in 6-well plates were co-transfected with pLKO1 mission MRE11 shRNA plasmids and helper plasmids psPax2 and pMD2.g. Six hours later, the medium was replaced with fresh DMEM without antibiotics. After 48 h, the recombinant lentivirus-containing culture medium was harvested and centrifuged at 15 min, 3,000 rpm, and 4°C to remove any floating cells and cell debris. The cleared media containing lentiviruses were at 1:3 and 1:10 v/v ratios, added to HCT116 cells and plated in a 12-well plate, and after 24 h; the media were replaced with the fresh cultivation medium; cell cultures containing integrated lentiviruses were selected by using 2 µg/ml of puromycin for 4–5 days. Transfected cells were then tested using

genomic PCR and Western blotting analysis for the genetic elimination/loss of expression of the *MRE11* gene.

5 Results

The result section will start with analysis of the expression of 3 MBNL paralogs and 7 control genes with transcription-variant-specific primers on a sample set of 108 patients. Descriptive statistics of enrolled patients are summarized in Table 7.

For the question whether the expression change in MBNL translate into the functional effect of the altered splicing of MBNL targets (FOXP1 and EPB41L3) or change in splicing in general (CD44), we analyzed selected alternatively spliced isoforms of EPB41L3, FOXP1 and CD44 genes that were previously confirmed to be regulated by MBNL or are known for their abundant alternative splicing.

5.1 Deregulation of MBNL and studied genes in tumor tissue

To analyze possible deregulated expression of the studied genes in CRC, we compared their expression in CRC tissue with the levels detected in paired non-tumorous mucosa by performing one sample t-tests on the respective -$\Delta\Delta$Ct values.

The analysis of MBNLs expression revealed significant decrease in the expression of all MBNL genes in tumor tissue compared to paired non-tumorous samples (MBNL1 p<0.001, MBNL2 p=0.047, MBNL3 p<0.001). It represents the first concise evidence that MBNL genes are deregulated in colorectal tumor cells.

Deregulation of gene expression was detected also for target gene panel, with CD44 (p<0.001), CD44v3 (p<0.001) and CD44v6 (p<0.001) showing higher expression in tumor tissue, and FOXP1com (p<0.001), FOXP1var (p=0.024), EPB41L3com (p<0.001), EPB41L3var (p<0.001) being upregulated in non-tumorous samples (Figure 7).

Expression of candidate genes in tumor tissue relative to healthy tissue
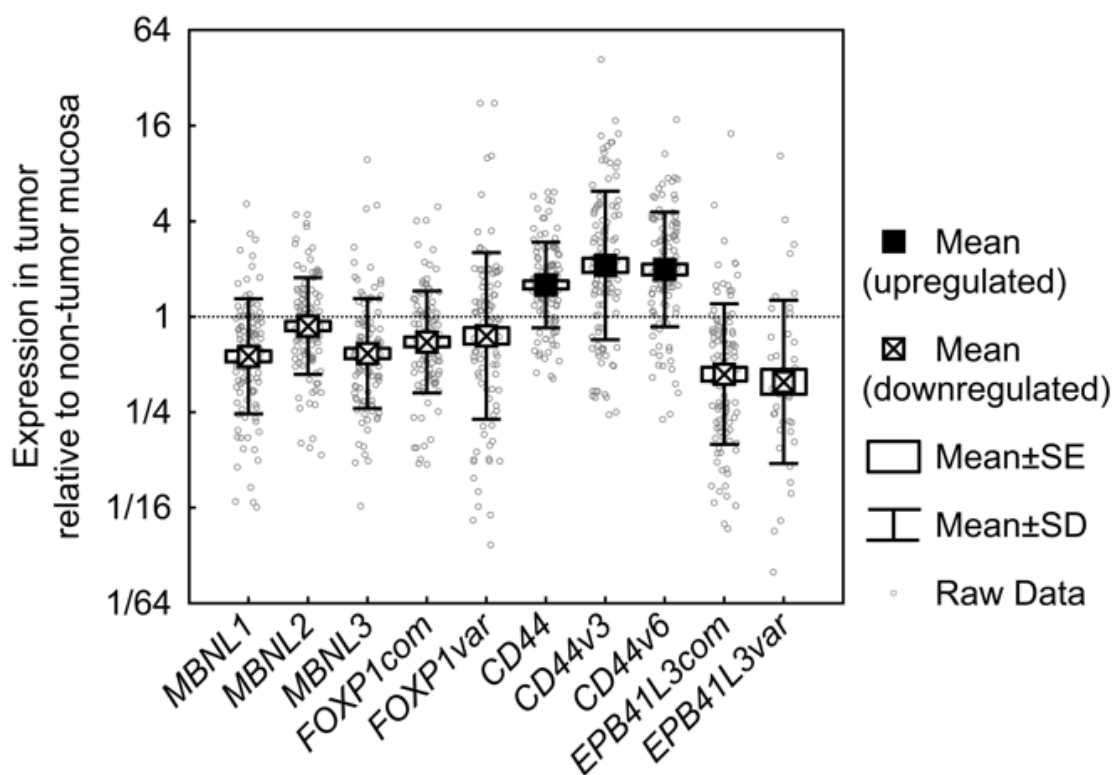

Figure 7. Expression of candidate genes in tumor tissue relative to healthy tissue. All genes show deregulation in tumor tissue.

Expression of studied genes *(–ΔΔ Ct* values) with corresponding statistical description

	Test of means vs reference constant							
	Mean	Stand. dev.	N	SEM	Reference constant	t	SV	p
-ΔΔCt MBNL1	-0.91420	1.346494	108	0.129566	0.00	-7.05587	107	0.000000
-ΔΔCt MBNL2	-0.19652	1.014078	108	0.097580	0.00	-2.01397	107	0.046522
-ΔΔCt MBNL3	-0.84128	1.354847	108	0.130370	0.00	-6.45301	107	0.000000
-ΔΔCt FOXP1com	-0.52659	1.066776	108	0.102651	0.00	-5.12996	107	0.000001
-ΔΔCt FOXP1var	-0.40139	1.746551	100	0.174655	0.00	-2.29818	99	0.023655
-ΔΔCt CD44	0.71911	1.056714	108	0.101682	0.00	7.07213	107	0.000000
-ΔΔCt CD44v3	1.07561	1.555374	108	0.149666	0.00	7.18674	107	0.000000
-ΔΔCt CD44v6	0.98878	1.199199	108	0.115393	0.00	8.56877	107	0.000000
-ΔΔCt EPB41L3com	-1.20303	1.471858	104	0.144328	0.00	-8.33541	103	0.000000
-ΔΔCt EPB41L3var	-1.30912	1.617895	106	0.157144	0.00	-8.33069	105	0.000000

Table 8. Expression of studied genes (–ΔΔ Ct values) with corresponding statistical description. The statistics were performed using one sample t test against zero constant.

5.2   Correlation between MBNL and target genes expression levels

Next, we analyzed correlation between MBNL expression and the target gene variant transcript ratios. Higher expression of MBNL1 and MBNL3 was significantly associated with increased variant transcript levels of FOXP1var (MBNL3 in tumor tissue p=0.046; MBNL1 in non-tumor tissues p=0.024; MBNL3 in non-tumorous tissues p=0.011). Similarly, MBNL3 overexpression was associated with increasing of variant transcript ratio of EPB41L3var in by non-malignant tissues (p=0.002). Otherwise, higher expression of MBNL2 in the tumor tissues was linked to lower transcript levels of EPB41L3var (p=0.011).

5.3   Expression of candidate genes according to T, N, M, G, and UICCC in tumor tissue relative to healthy tissue

5.3.1 In subgroups according to T

We analyzed three subgroups of patients with respect to T-staging (T2-T4; subgroup T1 was not analyzed due to small sample size n=3) (Figure 8). Accordingly, we adjusted significance level to p=0.05/3=0.0167. In general, direction of difference in genes expression between tumor tissue and adjacent mucosa tended to mirror the changes observed in pooled data. An exception was relative upregulation of FOXP1var in T4 subgroup, however, the effect was not recognized as significant.

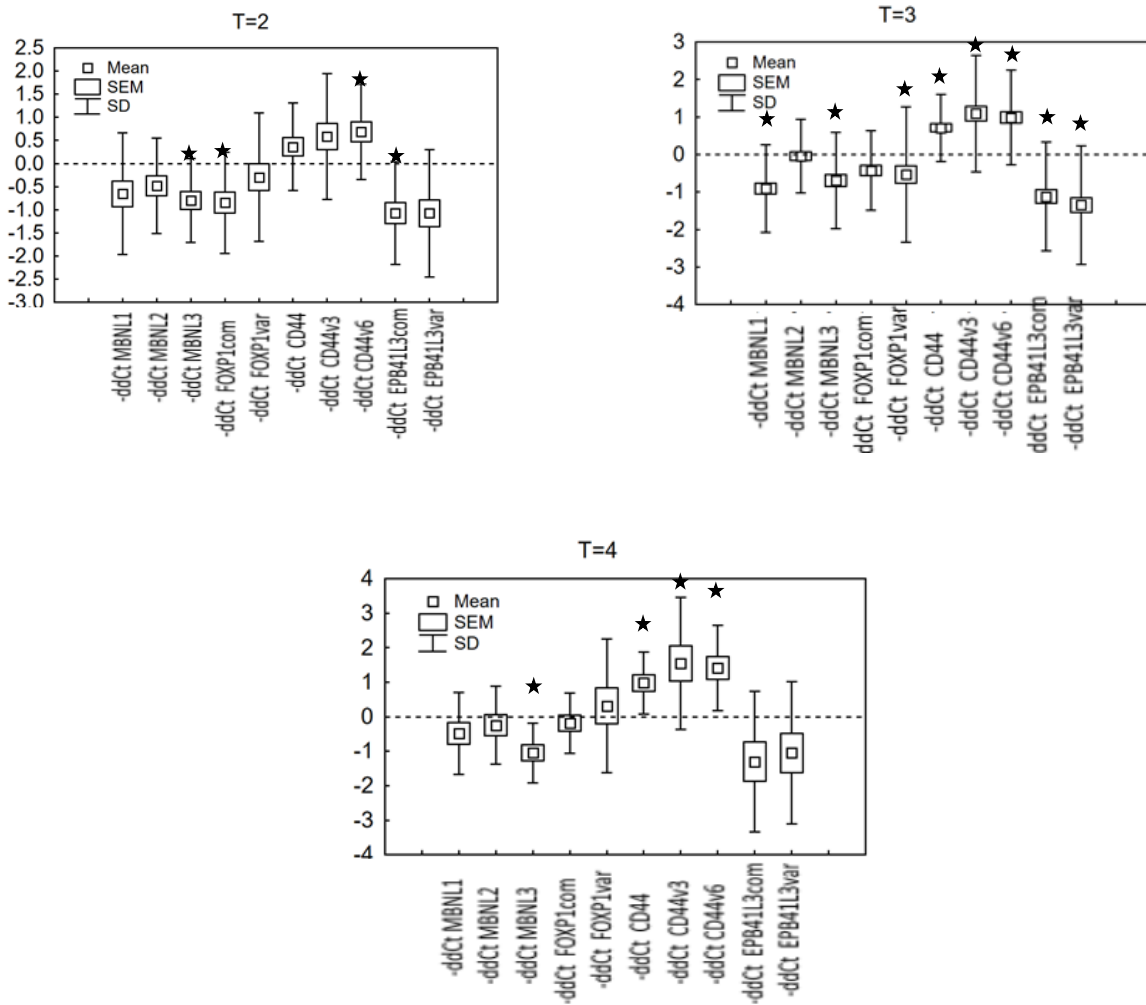Expression of studied genes according to T

Figure 8. Expression of studied genes according to T. Statistically significant deregulations denoted by asterisk mark.

5.3.2 In subgroups according to N

The patients were divided into 3 subgroups according to N-staging (N0-N1-N2, adjusted significance level p=0.0167) (Figure 9). The analysis performed on data within individual subgroups revealed differences in genes expression that were congruent with observations in pooled data. The downregulation of MBNL-2 and FOXP1var did not reach level of the significance in any of the subgroups.
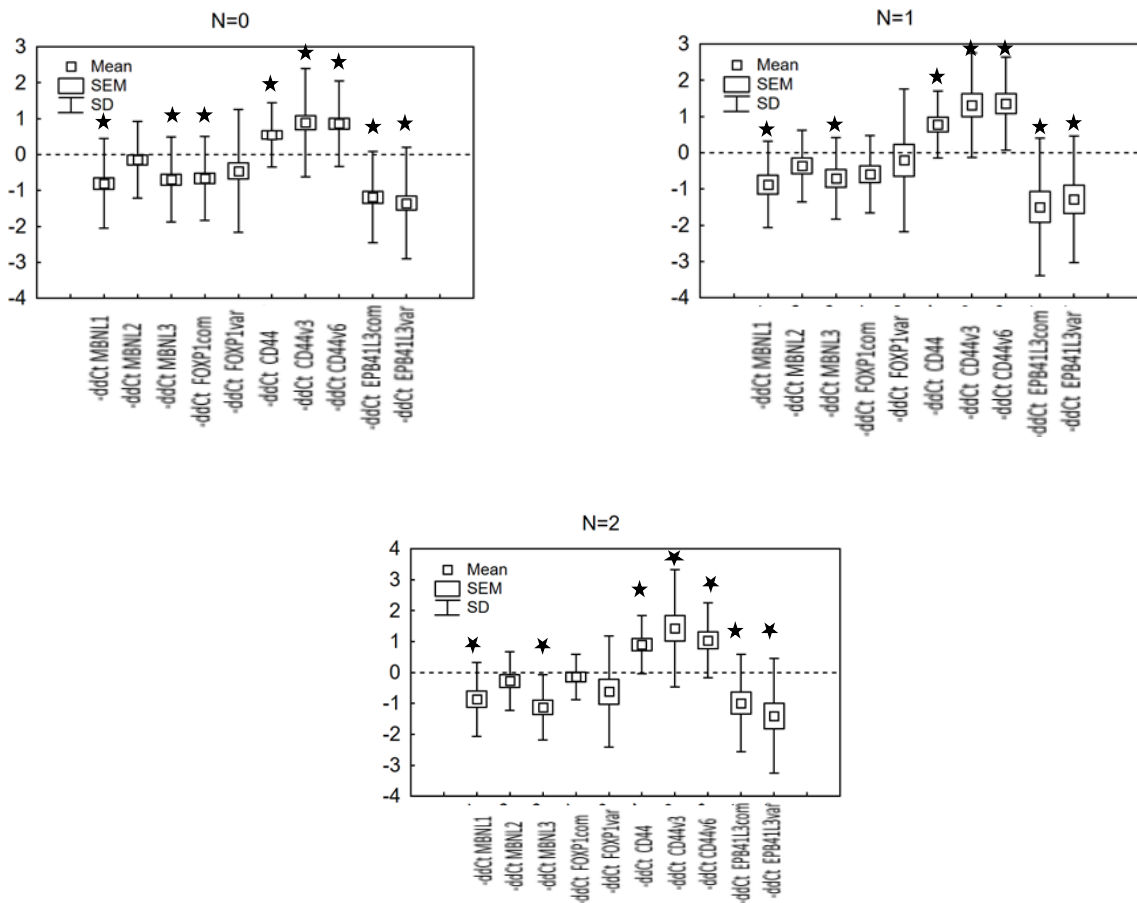
Expression of studied genes according to N

Figure 9. Expression of studied genes according to N. Statistically significant deregulations denoted by asterisk mark.

5.3.3 In subgroups according to M

We next analyzed data with respect to M staging (M0-M1, adjusted significance level p=0.025) (Figure 10). In M0 subgroup we detected significant deregulation of all the genes, consistently with the pooled data. In M1 subgroup there was no significant difference in expression of MBNL2, FOXP1com and FOXP1var.
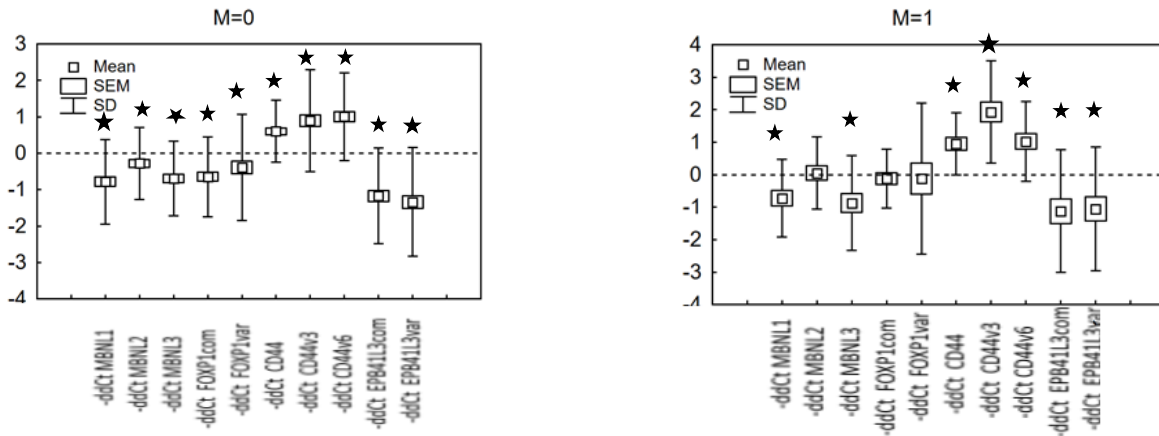
Expression of studied genes according M

Figure 10. Expression of studied genes according M. Statistically significant deregulations denotes by asterisk mark.

5.3.4 In subgroups according to G

In the analysis of subgroups with respect to tumor grading (G1-G3, adjusted significance level p=0.0167), we observed patterns similar to pooled data. Most of the detected deregulations were significant in G2 subgroup, likely reflecting the largest sample size. The MBNL-2 and FOXP1var did not display significant deregulation in any of the subgroups analyzed (Figure 11).
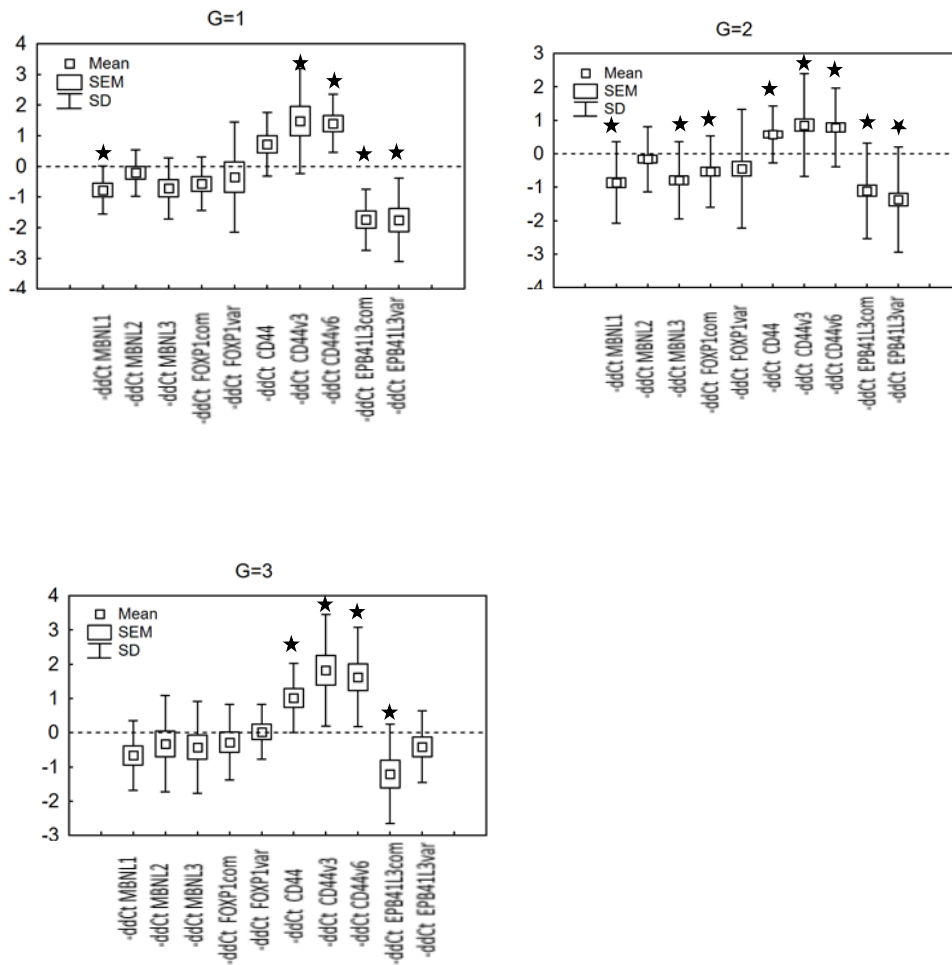
Expression of studied genes according to G

Figure 11. Expression of studied genes according to G. Statistically significant deregulations denotes by asterisk mark.

5.3.5 In subgroups according to UICC

We further analyzed data with respect to UICC staging (UICC1-UICC4, adjusted significance level p=0.013) (Figure 12). All the significant deregulations were consistent with analysis on the pooled dataset.
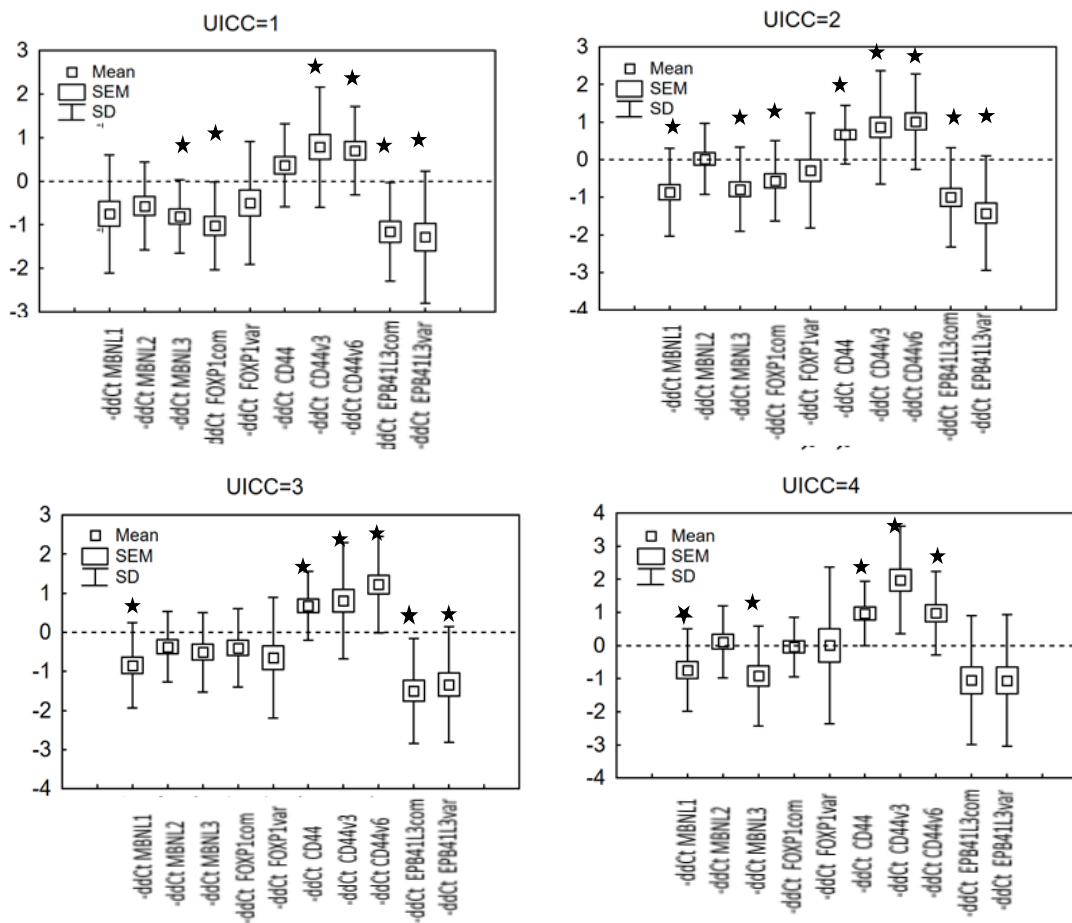
Expression of studied genes according UICC

Figure12. Expression of studied genes according UICC. Statistically significant deregulation denotes by asterisk mark.

5.4 Associations of genes expression with clinical factors (T, N, M, G, UICC)

Despite the deregulation in gene expression level, no correlation between MBNL1, 2 and 3 expression and clinical parameters was observed in our patients' set. We next analyzed associations of the target genes' expression with clinical parameters. The analysis detected positive trend for FOXP1com with both T- and N-staging advancement (T staging: Kendall correlation: τ =0.16, p= 0.014; N staging: Kendall correlation: τ =0.16, p=0.017) and CD44 with T-staging advancement (Kendall

correlation: т =0.15, p=0.020). Further, the comparison of expression between groups with and without distant metastasis at the time of surgery revealed significantly higher levels of FOXP1com expression (p=0.030) and significantly higher levels of relative expression of CD44v3 (p=0.020) in metastasis group. The overall UICC staging score was significantly correlated with expression of FOXP1com (Kendall correlation: т =0.23, p< 0.001), CD44 (Kendall correlation: т =0.15, p=0.019) and CD44v3 (Kendall correlation: т =0.16, p= 0.015). When considering tumor grading, we observed significantly non-uniform expression of CD44v6 across different grades (G1-G3) (one-way ANOVA: F=3.90, p=0.024) (Figure 13).

Expression of candidate genes according to T, N, M, G, and UICCC in tumor tissue relative to healthy tissue
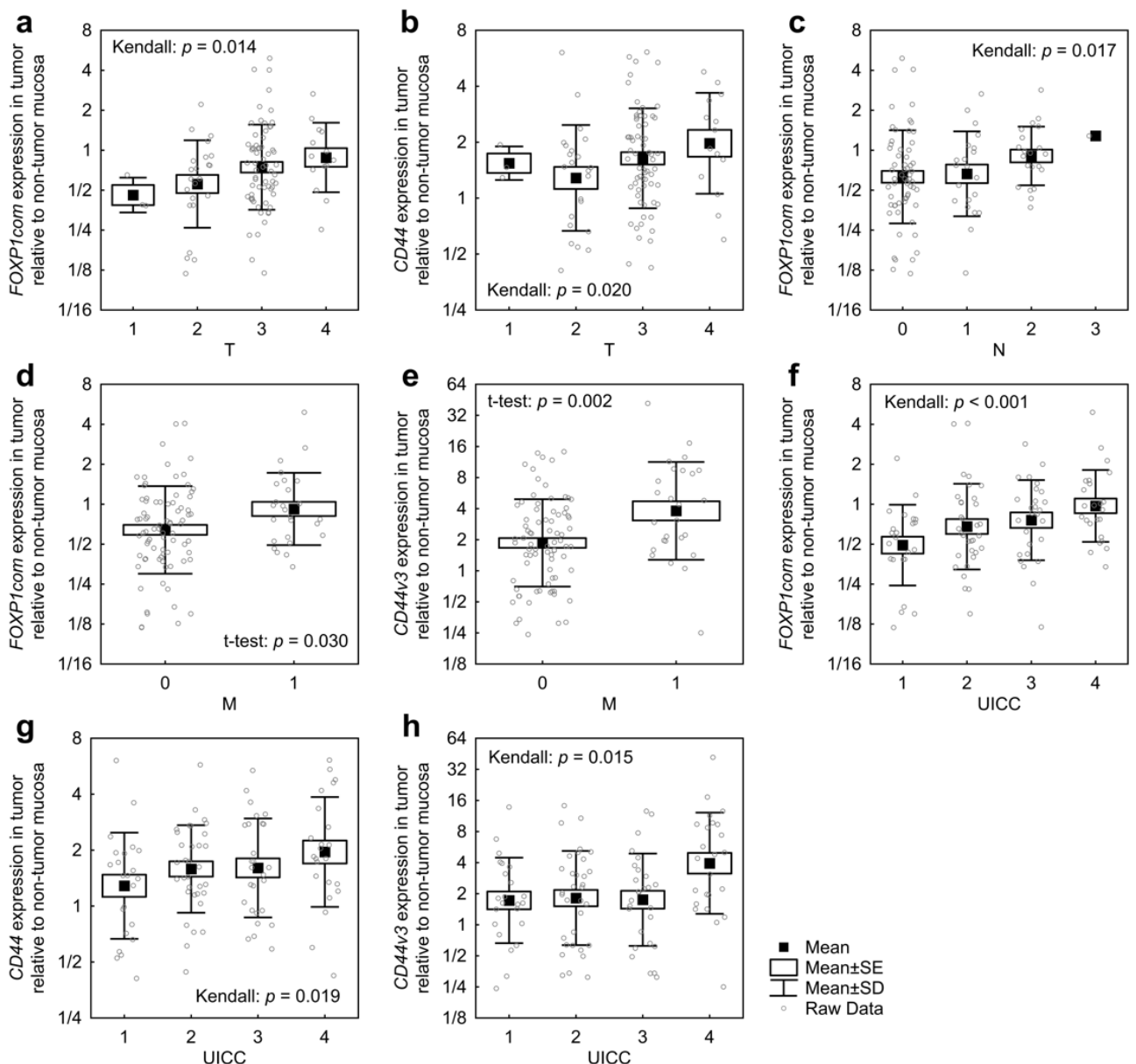
Figure 13. Expression of candidate genes according to T, N, M, G, and UICCC in tumor tissue relative to healthy tissue a) distribution of FOXP1com expression change according to T stage; b) distribution of CD44 expression change according to T stage; c) distribution of FOXP1com expression change according to N stage; d) distribution of FOXP1com expression change according to M stage; e) distribution of CD44v3 expression change according to M stage; f) distribution of FOXP1com expression change according to UICC score; g) distribution of CD44 expression change according to UICC score; h) distribution of CD44v3 expression change according to UICC score.

5.5   Survival analysis

Using univariable Cox proportional hazards model, proportional associations of the measured expression levels with overall survival were found for FOXPcom (p=0.002) with a hazard rate (HR) of 1.73 per unit of –ΔΔCt value (i.e. each doubling of the expression ratio between the tumor tissue and adjacent mucosa increases the risk of death 1.73-fold), FOXP1var (p<0.001, HR=1.50), and CD44v3 (p=0.023, HR=1.36). Recurrence-free survival was significantly influenced by FOXP1var (p=0.030, HR=1.26), CD44v3 (p=0.016, HR=1.33) and EPB41L3var EPB41L3com (p=0.025, HR=1.78). Subsequent automated determination of threshold values providing the best separation of groups confirmed these results (Figure 14,15).

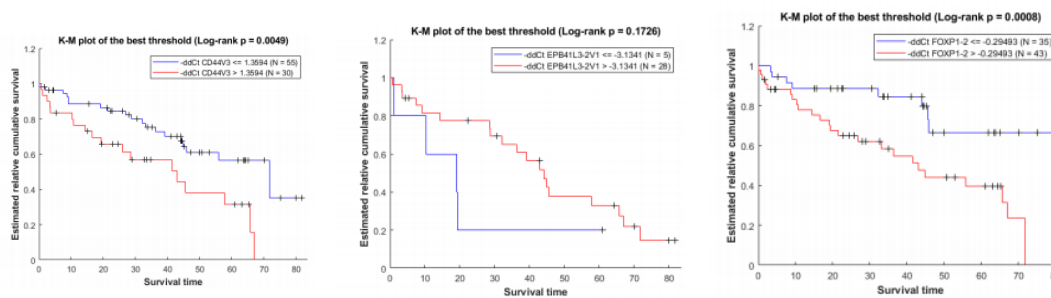Kaplan-Meier recurrence free survival curves for genes expression level

Fig 14. Kaplan-Meier recurrence free survival curves for genes expression level with respect for cut-off value a) CD44v3 b) EPB41L3var c) FOXP1var
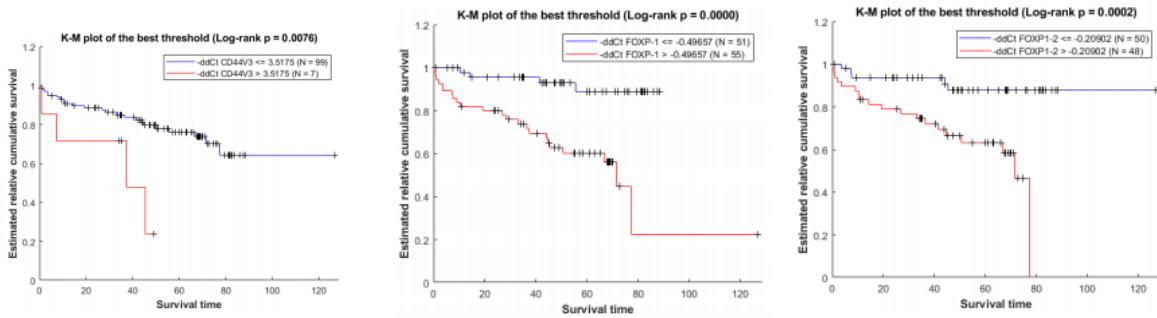
Fig15. Kaplan-Meier overall survival curves for genes expression level with respect for cut-off value a) CD44v3 b) FOXP1com c) FOXP1var

5.6 Study 2 Results

We detected significantly lower expression of both hOGG1 and MUTYH in CRC tissue than in adjacent non-malignant mucosa (p =0.0038 for MUTYH; p =0.0016 for hOGG1, Mann-Whitney U-test). Expression of hOGG1 was inversely associated with TNM staging (χ^2 test: p = 0.029). We did not detect significant relationship between MUTYH expression and TNM staging or particular MUTYH gene alterations.

Immunohistochemical analysis revealed highest hOGG1 expression in epithelial cells (adjacent to tumor and non-affected one), in contrast to relatively low level of hOGG1 in associated stroma. Notably, hOGG1 protein expression was significantly lower in CRC tissue compared to non-malignant epithelium, displaying approximately 50 % lower expression levels.

We next aimed to map presence of oncogenic mutations in KRAS gene. Overall, 51 KRAS mutations in 50 patients (29.5%) were detected. The most abundant substitutions were c.35 G>A (39.2%), followed by c.35 G>T (25.5%). We focused on G >T substitutions, potentially arising from oxidative damage and hOGG1/MUTYH repair system dysfunction. We observed 15 of these mutations in KRAS positive tumors (30%), corresponding to 7.8% of the tumour DNA samples across the whole dataset.

5.7   Study 3 Results

The TargetScan search revealed 187 miRNAs with MRE11 as a predicted target. These included 111 miRNAs with sufficient clinical data available. We then detected miR-140 as the miRNA with strongest association with progression-free survival (PFS).

We observed significantly decreased miR-140 levels in CRC tissue compared to paired non-malignant mucosa ($p<0.01$). The expression of MRE11 was non-significantly higher in tumour tissue ($p=0.11$). The low expression of miR-140 was associated with poor PFS ($p<0.05$), consistently with TCGA data.

We next applied miRNA mimics to induce miR-140 overexpression in colorectal cell lines. This resulted in decreased expression of both MRE11 protein and mRNA ($p<0.01$).

Further, the western blot analysis suggested that overexpression of MiR-140 in vitro led to increased levels of γH2AX protein, which is an established marker of DBS DNA damage.

We next studied effects of miR-140 overexpression in CRC cell line DLD1 on sensitivity to oxaliplatin treatment. Increasing miR-140 with miRNA mimics led to significantly decreased cell proliferation after oxaliplatin treatment ($p<0.05$), as well as decreased number of colonies ($p<0.05$), indicating decreased clonigenic potential. Overexpression of miR-140 did not enhance oxaliplatin-mediate supression of cell growth in shMRE11 cell line with suppressed MRE11 expression, suggesting that modulation of oxaliplatin sensitivity by miR-140 was mediated by MRE11.

6   Discussion

6.1   Study 1 Discussion

The existing studies on effect of altered expression of MBNL family mediated by alternative splicing factors in CRC are limited. The goal of this work was to establish the relationship between MBNL protein family expression status and to see if there is detectable deregulation of expression of the MBNL family in colorectal carcinoma. The first goal was to prove that MBNL protein family have aberrant expression patterns in CRC. The previous studies reported deregulation of individual MBNL paralogs in various cancers. It is of note that regulation of splicing events requires both MBNL1 and MBNL2 rather than an overall amount of non-specific MBNL protein. Recently, studies have shown unique functions of MBNL1 and MBNL2 in alternative splicing [214]. Knock-out of MBNL1 and MBNL2 results in different phenotypes in mouse. Splicing sensitive microarrays reveal that MBNL1 and MBNL2 regulate the alternative splicing of different pre-mRNAs. MBNL1 and MBNL2 could distinct their functions in different ways - they can also have completely various splicing functions, have dual functions where both proteins are required or they could have completely redundant function [247]. First, there are many splicing events that show mis-regulation when either MBNL1 or MBNL2 is knocked-out but they show normal splicing regulation when the other splicing factor is knocked-out instead [247,248]. For example, splicing patterns of specific proteins in mouse skeletal muscle (Clcn1(E7a) and Serca1(e22) and cardiac muscle (Tnnt2(e5) and Sorbs1(e25)) were shown to be regulated by MBNL1, while MBNL2 had no effect [249]. Comparably, in another study it was observed that in mouse embryonic fibroblast (*MEF*) cells, Mbnl1 and Lef1 are solely regulated by MBNL1 and MBNL2 has no effect [217]. While these studies restrict their analysis to specific cell types, it is likely that MBNL1 and MBNL2 similarly divide their functions across different tissues and conditions. This suggests that each of the MBNL paralogs requires specific attention when studying alternative splicing deregulation associated with cancer [246].

In the current study, we report deregulated expression of the MBNL proteins (MBNL1, MBNL2, MBNL3), which act as key regulators of alternative splicing and their increased activity promotes switch from embryonic to adult AS pattern, repress pluripotency

network and promotes differentiation. It is established that the expression profile of three MBNL paralogs and their correlated effect with the set of transcription factors might alter multiple splicing events. Our data show decreased MBNL expression in tumor tissue compared to matched non-malignant intestinal mucosa and simultaneous changes in the expression of splice-variants that are known to be regulated by MBNL. Marginally significant associations were observed between the expression of the studied genes and T, N, M, G and UICC clinical factors, with minute differences in gene expression among individual groups [246].

Observed downregulation of all three genes of the MBNL family in CRC tissue may promote less differentiated phenotype of the malignant cells. Previous studies proposed a role of the MBNL proteins in genesis of various tumor types [204]. It has been shown that MBNL1 inhibits invasiveness and migratory capacity of cancer cells, including metastasis of colorectal cells *in vitro*. Downregulation of MBNL3 was observed to be a crucial step in leukemia stem cell renewal [246]. On the other hand, MBNL2 is believed to be a driver of hypoxia adaptation, which is necessary for proliferation and migration of cancer cells in lung and breast cancer, and MBNL2 is thus considered a pro-oncogene in these cancer types. Function of the individual MBNL family members may therefore be very different and they can regulate particular alternative spicing patterns depending on the disease type and cellular context. In addition, function of MBNL can be regulated by their own alternative splicing, as was described for MBNL exon 7 in case of prostate cancer [246]. Although we detected downregulation of all members of the MBNL family, which is in line with their supposed tumor suppressor role, we observed neither an association of MBNL expression and T, N, M, G and UICC clinical factors in our data, nor any significant influence on patients' survival. This indicates MBNL downregulation as a significant marker of colorectal malignancy independent of disease progression, without direct manifestation of their downregulation at the level of clinicopathological parameters.

We further examined expressions of several other factors, whose alternative splicing was previously shown to be under the MBNL proteins' regulation and with implied role in cancer pathophysiology. The MBNL activity influences splicing of FOXP1, by promoting exclusion of exons that activate pluripotency genes and correspond to stem cell-like phenotype. Downregulation of MBNL might thus contribute to pluripotency network engagement in cancer cells by FOXP1-mediated effect. Overall, results of

previous studies on role of FOXP1 in cancer are controversial, suggesting both roles as an oncogene as well as tumor suppressor depending on the tumor type [240]. In our data, we observed significantly lower expression of FOXP1 in tumor cells, supporting previous findings in CRC. Unexpectedly, FOXP1 expression was positively associated with disease progression in terms of tumor invasion, nodes involvement and presence of distant metastasis. Moreover, higher expression of FOXP1 was associated with shortened overall survival [241].

These results seem to be in contrast with the previous study, where absence of FOXP1 was associated with decreased survival. Our results support the view of a more complex pattern of the FOXP1 factor involvement, possibly dependent on disease stage and on expression of particular isoforms. Under the condition of MBNL downregulation, expression of FOXP1 might correlate with less differentiated cell phenotype [229].

In addition, cancer tissue cells displayed enhanced expression of CD44 and its alternative splice variants CD44v3 and CD44v6. CD44 functions as an adhesion molecule in cell-cell and cell-matrix interactions and as a receptor in signal transduction, proposed to regulate tumor invasiveness and replication [250]. The observation of CD44 upregulation is consistent with a vast body of previous studies, however its role as a prognostic marker is not fully clear. Recent results of meta-analysis proposed its role as a negative prognostic marker [246]. This supports our observation that high CD44 expression was associated with tumor invasion and overall UICC staging.

As the CD44 is expressed in various alternatively spliced variants, we further focused on an analysis of CD44v3 and CD44v6, which have been proposed to play a direct role in carcinogenesis [251].

Downregulation of MBNL3 was shown to activate pluripotency network and CD44v3 overexpression in chronic myeloid leukemia [252]. A study with colon cancer stem lines proposed a role of CD44v3 in tumor invasiveness. CD44v6 is present in colorectal cancer stem cells, required for their ability to form metastasis and enhancing their resistance to chemotherapy both variants have been suggested as a prognostic marker in colorectal cancer, with a probably more established role of CD44v6 [253].

We show associations of alternative CD44 variants with several clinicopathological parameters of disease progression. The CD44v6 correlated with tumor grading

progression, while expression of CD44v3 was higher in patients with distant metastasis and correlated with overall UICC score. The higher expression of CD44v3 was also associated with worse survival rate. This highlights the role of CD44v3 next to CD44v6 as an important marker with potential prognostic value [254].

The upregulation of stem cell-like variants CD44v3 and CD44v6 is in line with the observed downregulation of MBNL and concurrent switch to stem cell phenotype with high potential towards pluripotency and invasiveness [250].

Another gene whose expression is associated with extensive alternative splicing is gene for cytoskeletal adaptor protein EPB41L3 that acts also as an important proliferation regulator. It has been demonstrated that alternative splicing of EPB41L3 is involved in differentiation of epithelial cells. The existent data is consistent with the role of EPB41L3 as a tumor suppressor [255,256]. Accordingly, we observed downregulation of EPB41L3 expression in colorectal cancer tissue.

6.2   Study 2 Discussion

We have studied activity of hOGG1 and MUTYH glycosylases, the important components of oxidative DNA damage repair system, in the context of sporadic CRC. We found decreased expression of both glycosylases in CRC tissue compared to non-malignant mucosa. The expression levels of hOGG1 were negatively associated with TNM score. The complementary activity of hOGG1 and MUTYH prevents G>T substitutions, particularly in KRAS oncogene, and we detected this mutation in subset of samples, corresponding to approximately one quarter of all the substitutions detected. The most frequent substitution was G>A, which might reflect various forms of oxidative damage [121].

6.3   Study 3 Discussion

We identified miR-140 as a miRNA targeting MRE11, a critical player in homolous recombination DNA repair, and found its decreased levels in CRC bioptic samples, which were associated with poor outcomes. Further *in vitro* analysis showed that

increasing miR-140 expression downregulates MRE11. We also found that increasing miR-140 enhances CRC cells' sensitivity to oxaliplatin and the effect is dependent on *MRE11* expression. These results indicate that miRNA might downregulate DNA repair factors, thereby potentiating efficiency of chemotherapy [156].

7 Conclusions

It is known that the expression profile of three MBNL paralogs and their correlated effect with the set of transcription factors might alter multiple splicing events. Our results data highlight the role of alternative splicing as an important mechanism in cancer pathogenesis. The alternative splicing thus, in addition to transcription changes, modulates genes expression to promote oncogenesis. Specifically, our data suggest a potential shift of the observed gene expressions towards stem-cell variants, associated with lower differentiation, pluripotency network engagement and enhanced invasiveness.

Altogether, in this study we have tried to advance our understanding of alternative splicing regulators as important markers of cancer progression and open the door for further systematic study of alternative splicing regulation in colorectal cancer and other malignancies.

In additional studies we examined selected factors of DNA damage repair in CRC. We found that CRC is associated with decreased hOGG1 and MUTYH glycosylases, which are crucially involved in repair of oxidative DNA damage. In another study we focused on factors involved in homologous recombination DNA repair. We showed that miR-140 downregulates MRE11, which enhances sensitivity of CRC cells to oxaliplatin *in vitro*. These results highlight involvement of DNA damage repair factors in CRC and their potential modulatory effect on CRC sensitivity to chemotherapy.

8   Literature

[1]   J.Ferlay, M. Colombet, I. Soerjomataram, C. Mathers, DM. Parkin, M. Piñeros, A. Znaor, F.Bray, Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods, Wiley-Liss Inc., 2019. doi:10.1002/ijc.31937.

[2]   A. Walther, E. Johnstone, C. Swanton, R. Midgley, I. Tomlinson, D. Kerr, Genetic prognostic and predictive markers in colorectal cancer, Nat. Rev. Cancer. 9 (2009) 489–499. doi:10.1038/nrc2645.

[3]   A.N. Archambault, Y.R. Su, J. Jeon, M. Thomas, Y. Lin, D. V. Conti, A.K. Win, L.C. Sakoda, I. Lansdorp-Vogelaar, E.F.P. Peterse, A.G. Zauber, D. Duggan, A.N. Holowatyj, J.R. Huyghe, H. Brenner, M. Cotterchio, S. Bézieau, S.L. Schmit, C.K. Edlund, M.C. Southey, R.J. MacInnis, P.T. Campbell, J. Chang-Claude, M.L. Slattery, A.T. Chan, A.D. Joshi, M. Song, Y. Cao, M.O. Woods, E. White, S.J. Weinstein, C.M. Ulrich, M. Hoffmeister, S.A. Bien, T.A. Harrison, J. Hampe, C.I. Li, C. Schafmayer, K. Offit, P.D. Pharoah, V. Moreno, A. Lindblom, A. Wolk, A.H. Wu, L. Li, M.J. Gunter, A. Gsur, T.O. Keku, R. Pearlman, D.T. Bishop, S. Castellví-Bel, L. Moreira, P. Vodicka, E. Kampman, G.G. Giles, D. Albanes, J.A. Baron, S.I. Berndt, S. Brezina, S. Buch, D.D. Buchanan, A. Trichopoulou, G. Severi, M.D. Chirlaque, M.J. Sánchez, D. Palli, T. Kühn, N. Murphy, A.J. Cross, A.N. Burnett-Hartman, S.J. Chanock, A. de la Chapelle, D.F. Easton, F. Elliott, D.R. English, E.J.M. Feskens, L.M. FitzGerald, P.J. Goodman, J.L. Hopper, T.J. Hudson, D.J. Hunter, E.J. Jacobs, C.E. Joshu, S. Küry, S.D. Markowitz, R.L. Milne, E.A. Platz, G. Rennert, H.S. Rennert, F.R. Schumacher, R.S. Sandler, D. Seminara, C.M. Tangen, S.N. Thibodeau, A.E. Toland, F.J.B. van Duijnhoven, K. Visvanathan, L. Vodickova, J.D. Potter, S. Männistö, K. Weigl, J. Figueiredo, V. Martín, S.C. Larsson, P.S. Parfrey, W.Y. Huang, H.J. Lenz, J.E. Castelao, M. Gago-Dominguez, V. Muñoz-Garzón, C. Mancao, C.A. Haiman, L.R. Wilkens, E. Siegel, E. Barry, B. Younghusband, B. Van Guelpen, S. Harlid, A. Zeleniuch-Jacquotte, P.S. Liang, M. Du, G. Casey, N.M. Lindor, L. Le Marchand, S.J. Gallinger, M.A. Jenkins, P.A. Newcomb, S.B. Gruber, R.E. Schoen, H. Hampel, D.A. Corley, L. Hsu, U. Peters, R.B. Hayes, Cumulative Burden of Colorectal Cancer-Associated Genetic Variants isMore Strongly Associated With Early-onset vs Late-onset Cancer, Gastroenterology. 158 (2020) 1274. doi:10.1053/J.GASTRO.2019.12.012.

[4]   Z. Zong, H. Li, C. Yi, H. Ying, Z. Zhu, H. Wang, Genome-Wide Profiling of Prognostic Alternative Splicing Signature in Colorectal Cancer, Front. Oncol. 8 (2018) 537.

doi:10.3389/fonc.2018.00537.

[5]   J. Zhang, Y. Deng, Y. Zuo, J. Wang, Y. Zhao, Analysis of Colorectal Cancer-
        Associated Alternative Splicing Based on Transcriptome, DNA Cell Biol. 39 (2020) 16–
        24. doi:10.1089/dna.2019.5111.

[6]   Colon Cancer Picture Image on MedicineNet.com, (n.d.).
        https://www.medicinenet.com/image-collection/colon_cancer_picture/picture.htm
        (accessed July 11, 2021).

[7]   I. Ewing, J.J. Hurley, E. Josephides, A. Millar, CURRICULUM BASED CLINICAL
        REVIEWS The molecular genetics of colorectal cancer, Frontline Gastroenterol. 5
        (2014) 26–30. doi:10.1136/flgastro-2013-100329.

[8]   P. Rawla, T. Sunkara, A. Barsouk, Epidemiology of colorectal cancer: Incidence,
        mortality, survival, and risk factors, Prz. Gastroenterol. 14 (2019) 89–103.
        doi:10.5114/pg.2018.81072.

[9]   B. Vogelstein, E.R. Fearon, S.R. Hamilton, S.E. Kern, A.C. Preisinger, M. Leppert,
        A.M. m. Smits, J.L. Bos, Genetic Alterations during Colorectal-Tumor Development, N.
        Engl. J. Med. 319 (1988) 525–532. doi:10.1056/NEJM198809013190901.

[10]   Y. Xi, P. Xu, Global colorectal cancer burden in 2020 and projections to 2040, Transl.
        Oncol. 14 (2021) 1936–5233. doi:10.1016/J.TRANON.2021.101174.

[11]   F. Bray, J. Ferlay, I. Soerjomataram, R.L. Siegel, L.A. Torre, A. Jemal, Global cancer
        statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36
        cancers in 185 countries, CA. Cancer J. Clin. 68 (2018) 394–424.
        doi:10.3322/caac.21492.

[12]   R.H.Bone, J.D.Cross, A.J.Dwyer, J.L.Fox, D.M.Hyams, K.Hassmiller Lich, T.A.
        Mackey, R.M. Miller, A.M. Fendrick, A Path to Improve Colorectal Cancer Screening
        Outcomes: Faculty Roundtable Evaluation of Cost-Effectiveness and Utility, n.d.

[13]   N.N. Keum, E. Giovannucci, Global burden of colorectal cancer: emerging trends, risk
        factors and prevention strategies, Nat. Rev. Gastroenterol. Hepatol. 16 (2019) 713–
        732. doi:10.1038/s41575-019-0189-8.

[14]   Cancer Today, (n.d.). https://gco.iarc.fr/today/home (accessed November 2, 2020).

[15]   Colorectal Cancer Awareness Month 2022 – IARC, (n.d.).
        https://www.iarc.who.int/featured-news/colorectal-cancer-awareness-month-2022/
        (accessed November 28, 2022).

[16]   G. Carioli, P. Bertuccio, P. Boffetta, F. Levi, C. La Vecchia, E. Negri, M. Malvezzi,
        European cancer mortality predictions for the year 2020 with a focus on prostate

cancer, Ann. Oncol. 31 (2020) 650–658. https://doi.org/10.1016/j.annonc.2020.02.009 (accessed September 28, 2020).

[17]   GLOBOCAN 2020: New Global Cancer Data | UICC, (n.d.). https://www.uicc.org/news/globocan-2020-new-global-cancer-data (accessed June 27, 2021).

[18]   C.M.Sturgeon, Tumour markers - an overview ScienceDirect Topics, (n.d.). https://www.sciencedirect.com/topics/medicine-and-dentistry/molecular-marker (accessed July 11, 2021).

[19]   K. Strimbu, J.A. Tavel, What are Biomarkers?, Curr. Opin. HIV AIDS. 5 (2010) 463. doi:10.1097/COH.0B013E32833ED177.

[20]   E. Koncina, S. Haan, S. Rauh, E. Letellier, Prognostic and predictive molecular biomarkers for colorectal cancer: Updates and challenges, Cancers (Basel). 12 (2020) 319. doi:10.3390/CANCERS12020319.

[21]   Y.H. Shaib, M. Rugge, D.Y. Graham, R.M. Genta, Management of gastric polyps: An endoscopy-based approach, Clin. Gastroenterol. Hepatol. 11 (2013) 1374–1384. doi:10.1016/j.cgh.2013.03.019.

[22]   C.R. Boland, A. Goel, S.G. Patel, The genetic and epigenetic landscape of early-onset colorectal cancer, Color. Cancer. 9 (2020) CRC23. doi:10.2217/crc-2020-0005.

[23]   G. Jung, E. Hernández-Illán, L. Moreira, F. Balaguer, A. Goel, Epigenetics of colorectal cancer: biomarker and therapeutic potential, Nat. Rev. Gastroenterol. Hepatol. 17 (2020) 111–130. doi:10.1038/s41575-019-0230-y.

[24]   H. Yamagishi, H. Kuroda, Y. Imai, H. Hiraishi, Molecular pathogenesis of sporadic colorectal cancers, Chin. J. Cancer. 35 (2016) 4. doi:10.1186/s40880-015-0066-y.

[25]   F. Caja, L. Vodickova, J. Kral, V. Vymetalkova, A. Naccarati, P. Vodicka, DNA Mismatch Repair Gene Variants in Sporadic Solid Cancers, Int. J. Mol. Sci. 21 (2020) 1–29. doi:10.3390/IJMS21155561.

[26]   B.A. Alves Martins, G.F. de Bulhões, I.N. Cavalcanti, M.M. Martins, P.G. de Oliveira, A.M.A. Martins, Biomarkers in Colorectal Cancer: The Role of Translational Proteomics Research, Front. Oncol. 9 (2019) 1284. doi:10.3389/FONC.2019.01284.

[27]   M. Marcuello, V. Vymetalkova, R.P.L. Neves, S. Duran-Sanchon, H.M. Vedeld, E. Tham, G. van Dalum, G. Flügen, V. Garcia-Barberan, R.J. Fijneman, A. Castells, P. Vodicka, G.E. Lind, N.H. Stoecklein, E. Heitzer, M. Gironella, Circulating biomarkers for early detection and clinical management of colorectal cancer, Mol. Aspects Med. 69 (2019) 107–122. doi:10.1016/J.MAM.2019.06.002.

[28]   T.Brody, Biological Marker - an overview ScienceDirect Topics, (n.d.).
https://www.sciencedirect.com/topics/medicine-and-dentistry/biological-marker
(accessed July 11, 2021).

[29]   L.Oyejide, O.R.Mendes, I.Mikaelian, Molecular Pathology: Applications in Nonclinical
Drug Development, (n.d.). https://www.sciencedirect.com/topics/agricultural-and-
biological-sciences/biomarkers (accessed July 11, 2021).

[30]   O.O. Ogunwobi, F. Mahmood, A. Akingboye, F.G. Haddad, R. Eid, H.R. Kourie, E.
Barouky, M. Ghosn, Biomarkers in Colorectal Cancer: Current Research and Future
Prospects, Int. J. Mol. Sci. 21 (2020) 1–20. doi:10.3390/IJMS21155311.

[31]   M.S. Pino, D.C. Chung, the chromosomal instability pathway in colon cancer,
Gastroenterology. 138 (2010). https://www.slideshare.net/spaliz/a-the-chromosomal-
instability-pathway-in-colon-cancer-184158841 (accessed October 3, 2017).

[32]   M. Vacante, A.M. Borzì, F. Basile, A. Biondi, Biomarkers in colorectal cancer: Current
clinical utility and future perspectives, 2018. /pmc/articles/PMC6288499/ (accessed
March 22, 2021).

[33]   N. Christou, J. Meyer, S. Popeskou, V. David, C. Toso, N. Buchs, E. Liot, J. Robert, F.
Ris, M. Mathonnet, Circulating Tumour Cells, Circulating Tumour DNA and Circulating
Tumour miRNA in Blood Assays in the Different Steps of Colorectal Cancer
Management, a Review of the Evidence in 2019, Biomed Res. Int. 2019 (2019).
doi:10.1155/2019/5953036.

[34]   G. Pellino, G. Gallo, P. Pallante, R. Capasso, A. De Stefano, I. Maretto, U. Malapelle,
S. Qiu, S. Nikolaou, A. Barina, G. Clerico, A. Reginelli, A. Giuliani, G. Sciaudone, C.
Kontovounisios, L. Brunese, M. Trompetto, F. Selvaggi, Noninvasive biomarkers of
colorectal cancer: Role in diagnosis and personalised treatment perspectives,
Gastroenterol. Res. Pract. 2018 (2018). doi:10.1155/2018/2397863.

[35]   L.J.W. Bosch, B. Carvalho, R.J.A. Fijneman, C.R. Jimenez, H.M. Pinedo, M. Van
Engeland, G.A. Meijer, Molecular tests for colorectal cancer screening, Clin. Colorectal
Cancer. 10 (2011) 8–23. doi:10.3816/CCC.2011.N.002.

[36]   N. Carlomagno, P. Incollingo, V. Tammaro, G. Peluso, N. Rupealta, G. Chiacchio,
M.L.S. Sotelo, G. Minieri, A. Pisani, E. Riccio, M. Sabbatini, U.M. Bracale, A.
Calogero, C.A. Dodaro, M. Santangelo, Diagnostic, Predictive, Prognostic, and
Therapeutic Molecular Biomarkers in Third Millennium: A Breakthrough in Gastric
Cancer, Biomed Res. Int. 2017 (2017). doi:10.1155/2017/7869802.

[37]   S. Spring, Diagnostic Biomarker - BEST (Biomarkers, EndpointS, and other Tools)
Resource - NCBI Bookshelf, (n.d.). https://www.ncbi.nlm.nih.gov/books/NBK402285/

(accessed July 11, 2021).

[38] T.Knösel, V.Emde, K.Schlüns, P.M.Schlag, M.Dietel, Cytokeratin profiles identify diagnostic signatures in colorectal cancer using multiplex analysis of tissue microarrays, Cell. Oncol. 28 (2006) 167–175. doi:10.1155/2006/354295.

[39] V.Baraka, H.Goike, K. W.Panaretakis, R.Einarsson, Clinical utility of cytokeratins as tumor markers, (n.d.). https://www.exbio.cz/about-exbio/blog/cytokeratins-markers-in-tumor-diagnostics (accessed July 11, 2021).

[40] S. Kummar, M. Fogarasi, A. Canova, A. Mota, T. Ciesielski, Cytokeratin 7 and 20 staining for the diagnosis of lung and colorectal adenocarcinoma, Br. J. Cancer. 86 (2002) 1884. doi:10.1038/SJ.BJC.6600326.

[41] B.D.W. Group, Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework, Clin. Pharmacol. Ther. 200 (2001) 89–95. doi:10.1067/mcp.2001.113989.

[42] C.C. Wu, T.W. Hsu, C.C. Yeh, H.B. Huang, The role of transcription factor caudal-related homeobox transcription factor 2 in colorectal cancer, Tzu-Chi Med. J. 32 (2020) 305. doi:10.4103/TCMJ.TCMJ_49_20.

[43] A.A. Bhat, A. Sharma, J. Pope, M. Krishnan, M.K. Washington, A.B. Singh, P. Dhawan, Caudal Homeobox Protein Cdx-2 Cooperates with Wnt Pathway to Regulate Claudin-1 Expression in Colon Cancer Cells, PLoS One. 7 (2012) e37174. doi:10.1371/JOURNAL.PONE.0037174.

[44] R. Pinto, R. Barros, I. Pereira-Castro, P. Mesquita, L.T. da Costa, E.P. Bennett, R. Almeida, L. David, CDX2 homeoprotein is involved in the regulation of ST6GalNAc-I gene in intestinal metaplasia, Lab. Investig. 2015 957. 95 (2015) 718–727. doi:10.1038/labinvest.2015.52.

[45] J. Brocato, M. Costa, SATB1 and 2 in colorectal cancer, Carcinogenesis. 36 (2015) 186–191. doi:10.1093/CARCIN/BGU322.

[46] B. Nodin, H. Johannesson, S. Wangefjord, D.P. O'Connor, K.E. Lindquist, M. Uhlén, K. Jirström, J. Eberhard, Molecular correlates and prognostic significance of SATB1 expression in colorectal cancer, Diagn. Pathol. 7 (2012). doi:10.1186/1746-1596-7-115.

[47] M.A. Mansour, T. Hyodo, K.A. Akter, T. Kokuryo, K. Uehara, M. Nagino, T. Senga, SATB1 and SATB2 play opposing roles in c-Myc expression and progression of colorectal cancer, Oncotarget. 7 (2016) 4993–5006. doi:10.18632/ONCOTARGET.6651.

[48]   A. Dragomir, M. De Wit, C. Johansson, M. Uhlen, F. Ponteń, The role of SATB2 as a diagnostic marker for tumors of colorectal origin: Results of a pathology-based clinical prospective study, Am. J. Clin. Pathol. 141 (2014) 630–638. doi:10.1309/AJCPWW2URZ9JKQJU.

[49]   Y.Li, Y. Liu, Y. Hu, L. Chen, J. Li, Special AT-rich sequence-binding protein 2 acts as a negative regulator of stemness in colorectal cancer cells, World J. Gastroenterol. 22 (2016) 8528–8539. doi:10.3748/WJG.V22.I38.8528.

[50]   N.S. Meagher, L. Wang, P.F. Rambau, M.P. Intermaggio, D.G. Huntsman, L.R. Wilkens, M.A. El-Bahrawy, R.B. Ness, K. Odunsi, H. Steed, E. Herpel, M.S. Anglesio, B. Zhang, N. Lambie, A.J. Swerdlow, J. Lubiński, R.A. Vierkant, E.L. Goode, U. Menon, A. Toloczko-Grabarek, O. Oszurek, S. Bilic, A. Talhouk, M. García-Closas, Q. Wang, A. Tan, R. Farrell, C.J. Kennedy, M. Jimenez-Linan, K. Sundfeldt, J.L. Etter, J. Menkiszak, M.T. Goodman, P. Klonowski, Y. Leung, S.J. Winham, K.B. Moysich, S. Behrens, T. Kluz, R.P. Edwards, J. Gronwald, F. Modugno, B.Y. Hernandez, C. Chow, L.E. Kelemen, G.L. Keeney, M.E. Carney, Y. Natanzon, G. Robertson, R. Sharma, S.A. Gayther, J. Alsop, H. Luk, C. Karpinskyj, I. Campbell, P. Sinn, A. Gentry-Maharaj, P. Coulson, J. Chang-Claude, M. Shah, M. Widschwendter, K. Tang, M.J. Schoemaker, J.M. Koziak, L.S. Cook, J.D. Brenton, F. Daley, B. Kristjansdottir, C. Mateoiu, M.C. Larson, P.R. Harnett, A. Jung, A. deFazio, K.L. Gorringe, P.D.P. Pharoah, P. Minoo, C. Stewart, O.F. Bathe, X. Gui, P. Cohen, S.J. Ramus, M. Köbel, A combination of the immunohistochemical markers CK7 and SATB2 is highly sensitive and specific for distinguishing primary ovarian mucinous tumors from colorectal and appendiceal metastases, Mod. Pathol. 2019 3212. 32 (2019) 1834–1846. doi:10.1038/s41379-019-0302-0.

[51]   S. Min-Cheng, Y. Ray-Hwang, L. Chiao-Ying, J. Yung-Ming, M. Su, RH. Yuan, C. Lin, Y. Jeng, Cadherin-17 is a useful diagnostic marker for adenocarcinomas of the digestive system, Mod. Pathol. 21 (2008) 1379–1386. doi:10.1038/MODPATHOL.2008.107.

[52]   E. Yakirevich, C. Magi-Galluzzi, Z. Grada, S. Lu, M.B. Resnick, S. Mangray, Cadherin 17 is a Sensitive and Specific Marker for Metanephric Adenoma, Am. J. Surg. Pathol. 39 (2015) 479. doi:10.1097/PAS.0000000000000401.

[53]   R.A. .Bartolomé, R.Barderas, S.Torres, M.J.Fernandez-Aceñero, M.Mendes, J.García-Foncillas, M.Lopez-Lucendo, J.I.Casal, Cadherin-17 interacts with α2β1 integrin to regulate cell proliferation and adhesion in colorectal cancer cells causing liver metastasis, (n.d.).

[54]   W. Baumgartner, Possible roles of LI-Cadherin in the formation and maintenance of the intestinal epithelial barrier, Tissue Barriers. 1 (2013) e23815. doi:10.4161/TISB.23815.

[55]   D. Tacha, D. Zhou, CDH17 is a Highly Specific and More Sensitive Marker Than CK20 and CDX2 in Colon Adenocarcinoma and in Stomach Adenocarcinoma www.biocare.net, (n.d.). www.biocare.net (accessed July 11, 2021).

[56]   O. Hyung-Hoon Hoon, J. Young-Eun Eun, Novel biomarkers for the diagnosis and prognosis of colorectal cancer, Intest. Res. 18 (2020) 168–183. doi:10.5217/IR.2019.00080.

[57]   L. Peng, J. Li, C. Meng, J. Li, D. Tang, F. Guan, P. Xu, T. Wei, Y. Li, Diagnostic Value of Telomerase Activity in Patients With Bladder Cancer: A Meta-Analysis of Diagnostic Test, Front. Oncol. 10 (2020) 2718. doi:10.3389/FONC.2020.570127.

[58]   C.M. Heaphy, A.K. Meeker, The potential utility of telomere-related markers for cancer diagnosis, Wiley-Blackwell, 2011. doi:10.1111/J.1582-4934.2011.01284.X.

[59]   M.A. Jafri, S.A. Ansari, M.H. Alqahtani, J.W. Shay, Roles of telomeres and telomerase in cancer, and advances in telomerase-targeted therapies, Genome Med. 8 (2016). doi:10.1186/S13073-016-0324-X.

[60]   E.Hiyama, K.Hiyama, Telomerase as tumor marker, (n.d.).

[61]   B. Heidenreich, R. Kumar, TERT promoter mutations in telomere biology, Mutat. Res. Rev. Mutat. Res. 771 (2017) 15–31. doi:10.1016/J.MRREV.2016.11.002.

[62]   B.B. Williams, N.C. Tebbutt, M. Buchert, T.L. Putoczki, K. Doggett, S. Bao, C.N. Johnstone, F. Masson, F. Hollande, A.W. Burgess, A.M. Scott, M. Ernst, J.K. Heath, Glycoprotein A33 deficiency: A new mouse model of impaired intestinal epithelial barrier function and inflammatory disease, DMM Dis. Model. Mech. 8 (2015) 805–815. doi:10.1242/DMM.019935.

[63]   B.B. Williams, N.C. Tebbutt, M. Buchert, T.L. Putoczki, K. Doggett, S. Bao, C.N. Johnstone, F. Masson, F. Hollande, A.W. Burgess, A.M. Scott, M. Ernst, J.K. Heath, Glycoprotein A33 deficiency: a new model of impaired intestinal epithelial barrier function and inflammatory disease, Dis. Model. Mech. (2015). doi:10.1242/DMM.019935:

[64]   P. Murer, L. Plüss, D. Neri, A novel human monoclonal antibody specific to the A33 glycoprotein recognizes colorectal cancer and inhibits metastasis, (n.d.).

[65]   V.I. Vanderstichele A, Busschaert P, Smeets D, Landolfo C, Van Nieuwenhuysen E, Leunen K, Neven P, Amant F, Mahner S, Braicu EI, Zeilinger R, Coosemans A,

Timmerman D, Lambrechts D, Chromosomal Instability in Cell-Free DNA as a Highly Specific Biomarker for Detection of Ovarian Cancer in Women with Adnexal Masses, Clin. Cancer Res. 23 (2017) 2223–2231. doi:10.1158/1078-0432.CCR-16-1078.

[66] N.S. Schütz E, Akbari MR, Beck J, Urnovitz H, Zhang WW, Bornemann-Kolatzki K, Mitchell WM, Nam RK, Chromosomal instability in cell-free DNA is a serum biomarker for prostate cancer, Clin. Chem. 61 (2015) 239–248. doi:10.1373/CLINCHEM.2014.226571.

[67] P. Aghagolzadeh, R. Radpour, New trends in molecular and cellular biomarker discovery for colorectal cancer, World J. Gastroenterol. 22 (2016) 5678. doi:10.3748/WJG.V22.I25.5678.

[68] A.J.X. Lee, D. Endesfelder, A.J. Rowan, A. Walther, N.J. Birkbak, P.A. Futreal, J. Downward, Z. Szallasi, I. P.M. Tomlinson, M. Howell, M. Kschischo, C. Swanton, Chromosomal instability confers intrinsic multidrug resistance, (n.d.).

[69] Z. Chen, C. Zhang, M. Zhang, B. Li, Y. Niu, L. Chen, J. Yang, S. Lu, J. Gao, L. Shen, Chromosomal instability of circulating tumor DNA reflect therapeutic responses in advanced gastric cancer, Cell Death Dis. 2019 1010. 10 (2019) 1–9. doi:10.1038/s41419-019-1907-4.

[70] A. Siskova, J. Kral, J. Drabova, K. Cervena, K. Tomasova, J. Jungwirth, T. Hucl, P. Kohout, S. Summerova, L. Vodickova, P. Vodicka, V. Vymetalkova, Discovery of Long Non-Coding RNA MALAT1 Amplification in Precancerous Colorectal Lesions, Int. J. Mol. Sci. 23 (2022). doi:10.3390/IJMS23147656.

[71] W.S. Tsang AH, Cheng KH, Wong AS, Ng SS, Ma BB, Chan CM, Tsui NB, Chan LW, Yung BY, Current and future molecular diagnostics in colorectal cancer and colorectal adenoma, World J. Gastroenterol. 20 (2014) 3847. doi:10.3748/WJG.V20.I14.3847.

[72] N. Hauptman, D. Glavač, Colorectal Cancer Blood-Based Biomarkers, Hindawi Limited, 2017. doi:10.1155/2017/2195361.

[73] H.T. Nguyen, H.-Q. Duong, The molecular characteristics of colorectal cancer: Implications for diagnosis and therapy, Oncol. Lett. 16 (2018) 9. doi:10.3892/OL.2018.8679.

[74] E. Chen, C.L. Cario, L. Leong, K. Lopez, C.P. Márquez, C. Chu, P.S. Li, E. Oropeza, I. Tenggara, J. Cowan, J.P. Simko, J.M. Chan, T. Friedlander, A.W. Wyatt, R. Aggarwal, P.L. Paris, P.R. Carroll, F. Feng, J.S. Witte, Cell-free DNA concentration and fragment size as a biomarker for prostate cancer, Sci. Reports 2021 111. 11 (2021) 1–9. doi:10.1038/s41598-021-84507-z.

[75] V.H. Agbor-Enoh S, Chan JL, Singh A, Tunc I, Gorham S, Zhu J, Pirooznia M,

Corcoran PC, Thomas ML, Lewis BGT, Jang MK, Ayares DL, Horvath KA, Mohiuddin MM, Circulating cell-free DNA as a biomarker of tissue injury: Assessment in a cardiac xenotransplantation model, J. Heart Lung Transplant. 37 (2018) 967–975. doi:10.1016/J.HEALUN.2018.04.009.

[76]   K. Strimbu, J. A. Tavel, What are biomarkers?, BMC Med. Genomics. 8 (2015). doi:10.1186/S12920-015-0109-X/METRICS.

[77]   J.P. Lopez, A. Diallo, C. Cruceanu, L.M. Fiori, S. Laboissiere, I. Guillet, J. Fontaine, J. Ragoussis, V. Benes, G. Turecki, C. Ernst, Biomarker discovery: quantification of microRNAs and other small non-coding RNAs using next generation sequencing, BMC Med. Genomics 2015 81. 8 (2015) 1–18. doi:10.1186/S12920-015-0109-X.

[78]   A. Finotti, E. Fabbri, I. Lampronti, J. Gasparello, M. Borgatti, R. Gambari, MicroRNAs and Long Non-coding RNAs in Genetic Diseases, Mol. Diagnosis Ther. 2019 232. 23 (2019) 155–171. doi:10.1007/S40291-018-0380-6.

[79]   C.N. Watson, A. Belli, V. Di Pietro, Small Non-coding RNAs: New Class of Biomarkers and Potential Therapeutic Targets in Neurodegenerative Disease, Front. Genet. 10 (2019) 364. doi:10.3389/FGENE.2019.00364.

[80]   M. Ratti, A. Lampis, M. Ghidini, M. Salati, M.B. Mirchev, N. Valeri, J.C. Hahne, MicroRNAs (miRNAs) and Long Non-Coding RNAs (lncRNAs) as New Tools for Cancer Therapy: First Steps from Bench to Bedside, Target. Oncol. 15 (2020) 261. doi:10.1007/S11523-020-00717-X.

[81]   F. Cao, Y. Hu, Z. Chen, W. Han, W. Lu, J. Xu, H. Ding, X. Shen, Circulating long noncoding RNAs as potential biomarkers for stomach cancer: a systematic review and meta-analysis, World J. Surg. Oncol. 2021 191. 19 (2021) 1–13. doi:10.1186/S12957-021-02194-6.

[82]   J. Fan, M. Saft, N. Sadanandan, B. Gonzales-Portillo, Y.J. Park, P.R. Sanberg, C. V. Borlongan, Y. Luo, LncRNAs Stand as Potent Biomarkers and Therapeutic Targets for Stroke, Front. Aging Neurosci. 12 (2020) 335. doi:10.3389/FNAGI.2020.594571.

[83]   S. Yuan, Y. Xiang, X. Guo, Y. Zhang, C. Li, W. Xie, N. Wu, L. Wu, T. Cai, X. Ma, Z. Yu, L. Bai, Y. Li, Circulating Long Noncoding RNAs Act as Diagnostic Biomarkers in Non-Small Cell Lung Cancer, Front. Oncol. 10 (2020) 2609. doi:10.3389/FONC.2020.537120.

[84]   Bolha L, Ravnik-Glavač M, Glavač D, Long Noncoding RNAs as Biomarkers in Cancer, Dis. Markers. 2017 (2017). doi:10.1155/2017/7243968.

[85]   T. Shi, G. Gao, Y. Cao, Long Noncoding RNAs as Novel Biomarkers Have a Promising Future in Cancer Diagnostics, Dis. Markers. 2016 (2016).

doi:10.1155/2016/9085195.

[86]   M. Svoboda, J. Slyskova, M. Schneiderova, P. Makovicky, L. Bielik, M. Levy, L. Lipska, B. Hemmelova, Z. Kala, M. Protivankova, O. Vycital, V. Liska, L. Schwarzova, L. Vodickova, P. Vodicka, HOTAIR long non-coding RNA is a negative prognostic factor not only in primary tumors, but also in the blood of colorectal cancer patients, Carcinogenesis. 35 (2014) 1510–1515. doi:10.1093/CARCIN/BGU055.

[87]   M. Barutaut, P. Fournier, W.F. Peacock, M.F. Evaristi, C. Caubère, A. Turkieh, F. Desmoulin, L.W.M. Eurlings, S. van Wijk, H.P.B. La Rocca, J. Butler, F. Koukoui, C. Dambrin, S. Mazeres, S. Le Page, C. Delmas, M. Galinier, C. Jung, F. Smih, P. Rouet, Insulin-like Growth Factor Binding Protein 2 predicts mortality risk in heart failure, Int. J. Cardiol. 300 (2020) 245–251. doi:10.1016/J.IJCARD.2019.09.032.

[88]   J. Yang, M. Griffiths, M.K. Nies, S. Brandal, R. Damico, D. Vaidya, X. Tao, C.E. Simpson, T.M. Kolb, S.C. Mathai, M.W. Pauciulo, W.C. Nichols, D.D. Ivy, E.D. Austin, P.M. Hassoun, A.D. Everett, Insulin-like growth factor binding protein-2: a new circulating indicator of pulmonary arterial hypertension severity and survival, BMC Med. 2020 181. 18 (2020) 1–11. doi:10.1186/S12916-020-01734-3.

[89]   Y. Lin, T. Jiang, K. Zhou, L. Xu, B. Chen, G. Li, X. Qiu, T. Jiang, W. Zhang, S.W. Song, Plasma IGFBP-2 levels predict clinical outcomes of patients with high-grade gliomas, Neuro. Oncol. 11 (2009) 468. doi:10.1215/15228517-2008-114.

[90]   G.J. Heald AH, Kaushal K, Siddals KW, Rudenski AS, Anderson SG, Insulin-like growth factor binding protein-2 (IGFBP-2) is a marker for the metabolic syndrome, Exp. Clin. Endocrinol. Diabetes. 114 (2006) 371–376. doi:10.1055/S-2006-924320.

[91]   J.M.Liou, C.T.Shun, J.T. Liang, H.M. Chiu, M.J. Chen, C.C. Chen, H.P. Wang, M.S. Wu, J.T. Lin, Plasma insulin-like growth factor-binding protein-2 levels as diagnostic and prognostic biomarker of colorectal cancer, J. Clin. Endocrinol. Metab. 95 (2010) 1717–1725. doi:10.1210/JC.2009-2668.

[92]   A. Sakuraba, N. Nemoto, N. Hibi, R. Ozaki, S. Tokunaga, O. Kikuchi, S. Minowa, T. Mitsui, M. Miura, D. Saito, M. Hayashida, J. Miyoshi, M. Matsuura, M. Yoneyama, H. Ohnishi, T. Hisamatsu, Extent of disease affects the usefulness of fecal biomarkers in ulcerative colitis, BMC Gastroenterol. 2021 211. 21 (2021) 1–8. doi:10.1186/S12876-021-01788-4.

[93]   M.A. Komor, L.J. Bosch, V.M. Coupé, C. Rausch, T. V Pham, S.R. Piersma, S. Mongera, C.J. Mulder, E. Dekker, E.J. Kuipers, M.A. van de Wiel, B. Carvalho, R.J. Fijneman, C.R. Jimenez, G.A. Meijer, M. de Wit, Proteins in stool as biomarkers for non-invasive detection of colorectal adenomas with high risk of progression, J. Pathol.

250 (2020) 288–298. doi:10.1002/PATH.5369.

[94] H.J. Galipeau, A. Caminero, W. Turpin, M. Bermudez-Brito, A. Santiago, J. Libertucci, M. Constante, J.A.R. Garay, G. Rueda, S. Armstrong, A. Clarizio, M.I. Smith, M.G. Surette, P. Bercik, P. Beck, C. Bernstein, K. Croitoru, L. Dieleman, B. Feagan, A. Griffiths, D. Guttman, K. Jacobson, G. Kaplan, D.O. Krause, K. Madsen, J. Marshall, P. Moayyedi, M. Ropeleski, E. Seidman, M. Silverberg, S. Snapper, A. Stadnyk, H. Steinhart, M. Surette, D. Turner, T. Walters, B. Vallance, G. Aumais, A. Bitton, M. Cino, J. Critch, L. Denson, C. Deslandres, W. El-Matary, H. Herfarth, P. Higgins, H. Huynh, J. Hyams, D. Mack, J. McGrath, A. Otley, R. Panancionne, K. Croitoru, E.F. Verdu, Novel Fecal Biomarkers That Precede Clinical Diagnosis of Ulcerative Colitis, Gastroenterology. 160 (2021) 1532–1545. doi:10.1053/J.GASTRO.2020.12.004.

[95] E.H. Schreuders, E.J. Grobbee, M.C.W. Spaander, E.J. Kuipers, Advances in Fecal Tests for Colorectal Cancer Screening, Curr. Treat. Options Gastroenterol. 14 (2016) 152. doi:10.1007/S11938-016-0076-0.

[96] A. Phalguni, H. Seaman, K. Routh, S. Halloran, S. Simpson, Tests detecting biomarkers for screening of colorectal cancer: What is on the horizon?, GMS Health Technol. Assess. 11 (2015) Doc01. doi:10.3205/HTA000122.

[97] A.G. Zauber, The Impact of Screening on Colorectal Cancer Mortality and Incidence – Has It Really Made a Difference?, Dig. Dis. Sci. 60 (2015) 681. doi:10.1007/S10620-015-3600-5.

[98] E.J. Grobbee, E. Wieten, B.E. Hansen, E.M. Stoop, T.R. de Wijkerslooth, I. Lansdorp-Vogelaar, P.M. Bossuyt, E. Dekker, E.J. Kuipers, M.C. Spaander, Fecal immunochemical test-based colorectal cancer screening: The gender dilemma, United Eur. Gastroenterol. J. 5 (2017) 448. doi:10.1177/2050640616659998.

[99] D.W. Ebner, J.B. Kisiel, Stool-Based Tests for Colorectal Cancer Screening: Performance Benchmarks Lead to High Expected Efficacy, Curr. Gastroenterol. Rep. 22 (2020). doi:10.1007/S11894-020-00770-6.

[100] T. Niedermaier, Y. Balavarca, H. Brenner, Stage-Specific Sensitivity of Fecal Immunochemical Tests for Detecting Colorectal Cancer: Systematic Review and Meta-Analysis, Am. J. Gastroenterol. 115 (2020) 56. doi:10.14309/AJG.0000000000000465.

[101] Y. Zhang, Q. Wu, L. Xu, H. Wang, X. Liu, S. Li, T. Hu, Y. Liu, Q. Peng, Z. Chen, X. Wu, J.-B. Fan, Sensitive detection of colorectal cancer in peripheral blood by a novel methylation assay, Clin. Epigenetics 2021 131. 13 (2021) 1–10. doi:10.1186/S13148-021-01076-8.

[102] S. Mo, H. Wang, L. Han, W. Xiang, W. Dai, P. Zhao, F. Pei, Z. Su, C. Ma, Q. Li, Z.

Wang, S. Cai, H. Wang, R. Liu, G. Cai, Fecal Multidimensional Assay for Non-Invasive Detection of Colorectal Cancer: Fecal Immunochemical Test, Stool DNA Mutation, Methylation, and Intestinal Bacteria Analysis, Front. Oncol. 11 (2021) 183. doi:10.3389/FONC.2021.643136.

[103] T. Pang, S.T. Leach, T. Katz, A.S. Day, C.Y. Ooi, Fecal Biomarkers of Intestinal Health and Disease in Children, Front. Pediatr. 2 (2014). doi:10.3389/FPED.2014.00006.

[104] G. Van Assche, Fecal Biomarkers for the Diagnosis and Management of Inflammatory Bowel Disease, Gastroenterol. Hepatol. (N. Y). 7 (2011) 396. /pmc/articles/PMC3151412/ (accessed July 11, 2021).

[105] M.H. Ebell, Accuracy of Fecal DNA and Fecal Immunochemical Test for Colorectal Cancer Detection, Am. Fam. Physician. 90 (2014) 326–331. http://goo.gl/3niWXb. (accessed July 11, 2021).

[106] W.T.D.M. DA Ahlquist, The stool DNA test is more accurate than the plasma septin 9 test in detecting colorectal neoplasia, Clin Gastroenterol Hepatol. 10 (2012) 272–277. doi:10.1016/j.cgh.2011.10.008.

[107] J.R. Jass, Classification of colorectal cancer based on correlation of clinical, morphological and molecular features, 2007.

[108] H. Brenner, M. Kloor, C.P. Pox, Colorectal cancer, Elsevier B.V., 2014. doi:10.1016/S0140-6736(13)61649-9.

[109] J. Wang, S. Liu, H. Wang, L. Zheng, C. Zhou, G. Li, R. Huang, H. Wang, C. Li, X. Fan, X. Fu, X. Wang, H. Guo, J. Guan, Y. Sun, X. Song, Z. Li, D. Mu, J. Sun, X. Liu, Y. Qi, F. Niu, C. Chen, X. Wu, X. Wang, X. Song, H. Zou, Robust performance of a novel stool DNA test of methylated SDC2 for colorectal cancer detection: a multicenter clinical study, Clin. Epigenetics 2020 121. 12 (2020) 1–12. doi:10.1186/S13148-020-00954-X.

[110] D.J.Robertson, T.F. Imperiale, Stool testing for colorectal cancer screening, Gastroenterology. 149 (2015) 192–206. doi:10.1053/j.gastro.2015.05.045.

[111] N.K. Osborn, D.A. Ahlquist, Stool screening for colorectal cancer: molecular approaches, Gastroenterology. 128 (2005) 1286–1293. doi:10.1053/j.gastro.2004.10.041.

[112] H. Rashid, B. Hossain, T. Siddiqua, M. Kabir, Z. Noor, M. Ahmed, R. Haque, Fecal MicroRNAs as Potential Biomarkers for Screening and Diagnosis of Intestinal Diseases, Front. Mol. Biosci. 7 (2020) 181. doi:10.3389/FMOLB.2020.00181.

[113] A. Link, F. Balaguer, Y. Shen, T. Nagasaka, J.J. Lozano, C.R. Boland, A. Goel, Fecal microRNAs as novel biomarkers for colon cancer screening, Cancer Epidemiol. Biomarkers Prev. 19 (2010) 1766. doi:10.1158/1055-9965.EPI-10-0027.

[114] T. Masuda, N. Hayashi, Y. Kuroda, S. Ito, H. Eguchi, K. Mimori, MicroRNAs as Biomarkers in Colorectal Cancer, Cancers (Basel). 9 (2017). doi:10.3390/CANCERS9090124.

[115] L. Cekaite, P.W. Eide, G.E. Lind, R.I. Skotheim, R.A. Lothe, MicroRNAs as growth regulators, their function and biomarker status in colorectal cancer, Oncotarget. 7 (2016) 6476. doi:10.18632/ONCOTARGET.6390.

[116] Y. Tung On, T. Ceen-Ming, E. K. Harriss, D. Benjamin, C. Polytarchou, Faecal microRNAs as a non-invasive tool in the diagnosis of colonic adenomas and colorectal cancer: A meta-analysis, Sci. Reports 2019 91. 9 (2019) 1–13. doi:10.1038/s41598-019-45570-9.

[117] N. Shirafkan, B. Mansoori, A. Mohammadi, N. Shomali, M. Ghasbi, B. Baradaran, MicroRNAs as novel biomarkers for colorectal cancer: New outlooks, Biomed. Pharmacother. 97 (2018) 1319–1330. doi:10.1016/J.BIOPHA.2017.11.046.

[118] B.A. Alves Martins, G.F. de Bulhões, I.N. Cavalcanti, M.M. Martins, P.G. de Oliveira, A.M.A. Martins, Biomarkers in Colorectal Cancer: The Role of Translational Proteomics Research, Front. Oncol. 9 (2019) 1284. doi:10.3389/fonc.2019.01284.

[119] P. Vodicka, M. Urbanova, P. Makovicky, K. Tomasova, M. Kroupa, R. Stetina, A. Opattova, K. Kostovcikova, A. Siskova, M. Schneiderova, V. Vymetalkova, L. Vodickova, Oxidative damage in sporadic colorectal cancer: Molecular mapping of base excision repair glycosylases in colorectal cancer patients, Int. J. Mol. Sci. 21 (2020). doi:10.3390/ijms21072473.

[120] M.R. Murphy, F.E. Kleiman, Connections between 3′ end processing and DNA damage response: Ten years later, Wiley Interdiscip. Rev. RNA. 11 (2020) e1571. doi:10.1002/WRNA.1571.

[121] M.J. Kavec, M. Urbanova, P. Makovicky, A. Opattová, K. Tomasova, M. Kroupa, K. Kostovcikova, A. Siskova, N. Navvabi, M. Schneiderova, V. Vymetalkova, L. Vodickova, P. Vodicka, Oxidative Damage in Sporadic Colorectal Cancer: Molecular Mapping of Base Excision Repair Glycosylases MUTYH and hOGG1 in Colorectal Cancer Patients, Int. J. Mol. Sci. 23 (2022) 5704. doi:10.3390/ijms23105704.

[122] P. Vodicka, S. Vodenkova, T. Buchler, L. Vodickova, DNA repair capacity and response to treatment of colon cancer, Pharmacogenomics. 20 (2019) 1225–1233. doi:10.2217/pgs-2019-0070.

[123] S. Vodenkova, K. Jiraskova, M. Urbanova, M. Kroupa, J. Slyskova, M. Schneiderova, M. Levy, T. Buchler, V. Liska, L. Vodickova, V. Vymetalkova, A. Collins, A. Opattova, P. Vodicka, Base excision repair capacity as a determinant of prognosis and therapy response in colon cancer patients, DNA Repair (Amst). 72 (2018) 77–85. doi:10.1016/j.dnarep.2018.09.006.

[124] A.M. Rachiglio, A. Sacco, L. Forgione, C. Esposito, N. Chicchinelli, N. Normanno, Colorectal cancer genomic biomarkers in the clinical management of patients with metastatic colorectal carcinoma, Explor. Target. Anti-Tumor Ther. 1 (2020) 53–70. doi:10.37349/ETAT.2020.00004.

[125] J.N. Patel, M.K. Fong, M. Jagosky, Colorectal Cancer Biomarkers in the Era of Personalized Medicine, J. Pers. Med. 9 (2019). doi:10.3390/JPM9010003.

[126] W. Chen, B.J. Swanson, W.L. Frankel, Molecular genetics of microsatellite-unstable colorectal cancer for pathologists, Diagn. Pathol. 12 (2017). doi:10.1186/S13000-017-0613-8.

[127] C.R. Boland, A. Goel, Microsatellite Instability in Colorectal Cancer, NIH Public Access, 2010. doi:10.1053/J.GASTRO.2009.12.064.

[128] B. Maraqa, G. Al-Shbool, O. Abu-Shawer, M. Souleiman, O. Alshakhatreh, A. Al-Omari, H. Abdelkhaleq, A. Taqash, M. Al-Hussaini, Frequency of Mismatch Repair Protein (MMRP) Deficiency among Young Jordanians Diagnosed with Colorectal Carcinoma (CRC), Gastroenterol. Res. Pract. 2020 (2020). doi:10.1155/2020/5632984.

[129] F.A. Sinicrope, N.R. Foster, S.N. Thibodeau, S. Marsoni, G. Monges, R. Labianca, G.P. Kim, G. Yothers, C. Allegra, M.J. Moore, S. Gallinger, D.J. Sargent, DNA mismatch repair status and colon cancer recurrence and survival in clinical trials of 5-fluorouracil-based adjuvant therapy, J. Natl. Cancer Inst. 103 (2011) 863–875. doi:10.1093/JNCI/DJR153.

[130] M-C.Etienne-Grimaldi., A.Mahamat, M.Chazal, P.Laurent-Puig, S.Olschwang, M-P. Gaub., J.L.Formento, P.Formento, A.Sudaka, V.Boige, A.Abderrahim-Ferkoune, D.Benchimol, T.André, S.Houry, J-L.Faucheron, C.PLetoublon, F.N.Gilly, J-R. Delpero, P.Lasser, B. Pradere, D.Pezet, F.Penault-Llorca, G.Milano, Molecular patterns in deficient mismatch repair colorectal tumours: results from a French prospective multicentric biological and genetic study, Br. J. Cancer 2014 11011. 110 (2014) 2728–2737. doi:10.1038/bjc.2014.213.

[131] M.Kloor, M.von Knebel Doeberitz, The Immune Biology of Microsatellite-Unstable Cancer, Trends in Cancer. 2 (2016) 121–133. doi:10.1016/J.TRECAN.2016.02.004.

[132] E.N. Mojarad, P.J. Kuppen, H.A. Aghdaei, M.R. Zali, The CpG island methylator phenotype (CIMP) in colorectal cancer, Gastroenterol. Hepatol. From Bed to Bench. 6 (2013) 120. /pmc/articles/PMC4017514/ (accessed July 12, 2021).

[133] M. Toyota, N. Ahuja, M. Ohe-Toyota, J.G. Herman, S.B. Baylin, J.-P.J. Issa, CpG island methylator phenotype in colorectal cancer, Proc. Natl. Acad. Sci. U. S. A. 96 (1999) 8681. doi:10.1073/PNAS.96.15.8681.

[134] L. Zong, Y. Seto, CpG island methylator phenotype, Helicobacter pylori, Epstein-Barr virus, and microsatellite instability and prognosis in gastric cancer: A systematic review and meta-analysis, PLoS One. 9 (2014). doi:10.1371/JOURNAL.PONE.0086097.

[135] J.H. Kim, G.H. Kang, Molecular and prognostic heterogeneity of microsatellite-unstable colorectal cancer, World J. Gastroenterol. 20 (2014) 4230–4243. doi:10.3748/WJG.V20.I15.4230.

[136] C.H. Kim, J.W. Huh, H.R. Kim, Y.J. Kim, CpG island methylator phenotype is an independent predictor of survival after curative resection for colorectal cancer: A prospective cohort study, J. Gastroenterol. Hepatol. 32 (2017) 1469–1474. doi:10.1111/JGH.13734.

[137] J.M. Bae, J.H. Kim, Y. Kwak, D.-W. Lee, Y. Cha, X. Wen, T.H. Lee, N.-Y. Cho, S.-Y. Jeong, K.J. Park, S.W. Han, H.S. Lee, T.-Y. Kim, G.H. Kang, Distinct clinical outcomes of two CIMP-positive colorectal cancer subtypes based on a revised CIMP classification system, Br. J. Cancer. 116 (2017) 1012. doi:10.1038/BJC.2017.52.

[138] O. Menyhárt, Á. Nagy, B. Győrffy, Determining consistent prognostic biomarkers of overall survival and vascular invasion in hepatocellular carcinoma, R. Soc. Open Sci. 5 (2018). doi:10.1098/RSOS.181006.

[139] S. Mallett, A. Timmer, W. Sauerbrei, D.G. Altman, Reporting of prognostic studies of tumour markers: a review of published articles in relation to REMARK guidelines, Br. J. Cancer. 102 (2010) 173–180. doi:10.1038/SJ.BJC.6605462).

[140] N.N. Basu, S. Ingham, J. Hodson, F. Lalloo, M. Bulman, A. Howell, D.G. Evans, Risk of contralateral breast cancer in BRCA1 and BRCA2 mutation carriers: a 30-year semi-prospective analysis, Fam. Cancer. 14 (2015) 531–538. doi:10.1007/S10689-015-9825-9.

[141] T.H. Chen, S.W. Chang, C.C. Huang, K.L. Wang, K.T. Yeh, C.N. Liu, H.Lee, C.C.Lin, Y.W.Cheng, The prognostic significance of APC gene mutation and miR-21 expression in advanced-stage colorectal cancer, Colorectal Dis. 15 (2013) 1367–1374. doi:10.1111/CODI.12318.

[142] Li. XL, J.Zhou, ZR.Chen, WJ.Chng, p53 mutations in colorectal cancer- molecular pathogenesis and pharmacological reactivation, World J. Gastroenterol. 21 (2015) 84. doi:10.3748/WJG.V21.I1.84.

[143] A. Malki, R.A. ElRuz, I. Gupta, A. Allouch, S. Vranic, A.-E. Al Moustafa, Molecular Mechanisms of Colon Cancer Progression and Metastasis: Recent Insights and Advancements, Int. J. Mol. Sci. 22 (2021) 1–24. doi:10.3390/IJMS22010130.

[144] J.Chen, F.Wu, J.H. HL.Pei ,WD. Gu , ZH.Ning , YJ.Shao, Analysis of the correlation between P53 and Cox-2 expression and prognosis in esophageal cancer, (n.d.).

[145] P. Zhan, Y.-N. Ji, L.-K. Yu, TP53 mutation is associated with a poor outcome for patients with hepatocellular carcinoma: evidence from a meta-analysis, Hepatobiliary Surg. Nutr. 2 (2013) 260. doi:10.3978/J.ISSN.2304-3881.2013.07.06.

[146] I. Wasserman, L.H. Lee, S. Ogino, M.R. Marco, C. Wu, X. Chen, J. Datta, E. Sadot, B. Szeglin, J.G. Guillem, P.B. Paty, M.R. Weiser, G.M. Nash, L. Saltz, A. Barlas, K. Manova-Todorova, S.P.B. Uppada, A.E. Elghouayel, P. Ntiamoah, J.N. Glickman, T. Hamada, K. Kosumi, K. Inamura, A.T. Chan, R. Nishihara, A. Cercek, K. Ganesh, N.E. Kemeny, P. Dhawan, R. Yaeger, C.L. Sawyers, J. Garcia-Aguilar, M. Giannakis, J. Shia, J.J. Smith, SMAD4 Loss in Colorectal Cancer Patients Correlates with Recurrence, Loss of Immune Infiltrate, and Chemoresistance, Clin. Cancer Res. 25 (2019) 1948–1956. doi:10.1158/1078-0432.CCR-18-1726.

[147] L.Levy, C.S.Hill, Smad4 dependency defines two classes of transforming growth factor {beta} (TGF-{beta}) target genes and distinguishes TGF-{beta}-induced epithelial-mesenchymal transition from its antiproliferative and migratory responses, Mol. Cell. Biol. 25 (2005) 8108–8125. doi:10.1128/MCB.25.18.8108-8125.2005.

[148] A.Z. Becerra, C.P. Probst, M.A. Tejani, C.T. Aquina, M.G. González, B.J. Hensley, K. Noyes, J.R. Monson, F.J. Fleming, Evaluating the Prognostic Role of Elevated Preoperative Carcinoembryonic Antigen Levels in Colon Cancer Patients: Results from the National Cancer Database, Ann. Surg. Oncol. 23 (2016) 1554–1561. doi:10.1245/S10434-015-5014-1.

[149] P.L. Tsai, W.J. Su, W.H. Leung, C.T. Lai, C.K. Liu, Neutrophil-lymphocyte ratio and CEA level as prognostic and predictive factors in colorectal cancer: A systematic review and meta-analysis, J. Cancer Res. Ther. 12 (2016) 582–589. doi:10.4103/0973-1482.144356.

[150] H. Kim, H. Il Jung, S.H. Kwon, S.H. Bae, H.C. Kim, M.-J. Baek, M.S. Lee, Preoperative neutrophil-lymphocyte ratio and CEA is associated with poor prognosis in patients with synchronous colorectal cancer liver metastasis, Ann. Surg. Treat. Res. 96 (2019) 191.

doi:10.4174/ASTR.2019.96.4.191.

[151] Y.Yeh, P.Lin, C.Lee, et al., Treatment monitoring of colorectal cancer by integrated analysis of plasma concentration and sequencing of circulating tumor DNA, Mol. Cancer 2020 191. 19 (2020) 1–6. doi:10.1186/S12943-020-01273-8.

[152] D. Fernandez-Garcia, A. Hills, K. Page, R.K. Hastings, B. Toghill, K.S. Goddard, C. Ion, O. Ogle, A.R. Boydell, K. Gleason, M. Rutherford, A. Lim, D.S. Guttery, R.C. Coombes, J.A. Shaw, Plasma cell-free DNA (cfDNA) as a predictive and prognostic marker in patients with metastatic breast cancer, Breast Cancer Res. 2019 211. 21 (2019) 1–13. doi:10.1186/S13058-019-1235-8.

[153] A.J. Bronkhorst, V. Ungerer, S. Holdenrieder, The emerging role of cell-free DNA as a molecular marker for cancer management, Biomol. Detect. Quantif. 17 (2019). doi:10.1016/J.BDQ.2019.100087.

[154] K. Ihara, S. Yamaguchi, N. Ueno, Y. Tani, Y. Shida, H. Ogata, Y. Domeki, K. Okamoto, M. Nakajima, K. Sasaki, T. Tsuchioka, H. Mitomi, H. Kato, Expression of DNA double-strand break repair proteins predicts the response and prognosis of colorectal cancer patients undergoing oxaliplatin-based chemotherapy, Oncol. Rep. 35 (2016) 1349–1355. doi:10.3892/or.2015.4488.

[155] A. Naccarati, F. Rosa, V. Vymetalkova, E. Barone, K. Jiraskova, C. Di Gaetano, J. Novotny, M. Levy, L. Vodickova, F. Gemignani, T. Buchler, S. Landi, P. Vodicka, B. Pardini, Double-strand break repair and colorectal cancer: Gene variants within 3' UTRs and microRNAs binding as modulators of cancer risk and clinical outcome, Oncotarget. 7 (2016) 23156–23169. doi:10.18632/oncotarget.6804.

[156] J. Horak, A. Dolnikova, O. Cumaogullari, A. Cumova, N. Navvabi, L. Vodickova, M. Levy, M. Schneiderova, V. Liska, L. Andera, P. Vodicka, A. Opattova, MiR-140 leads to MRE11 downregulation and ameliorates oxaliplatin treatment and therapy response in colorectal cancer patients, Front. Oncol. 12 (2022). doi:10.3389/FONC.2022.959407/FULL.

[157] L. Hui-Yan, X. Rui-Hua, Predictive and prognostic biomarkers with therapeutic targets in advanced colorectal cancer, World J. Gastroenterol. 20 (2014) 3858. doi:10.3748/WJG.V20.I14.3858.

[158] C. Lo Nigro, V. Ricci, D. Vivenza, C. Granetto, T. Fabozzi, E. Miraglio, M.C. Merlano, Prognostic and predictive biomarkers in metastatic colorectal cancer anti-EGFR therapy, World J. Gastroenterol. 22 (2016) 6944. doi:10.3748/WJG.V22.I30.6944.

[159] A.Rowland, MM. Dias, MD. Wiese, G. Kichenadasse, RA. McKinnon, CS. Karapetis, MJ. Sorich, Meta-analysis of BRAF mutation as a predictive biomarker of benefit from

anti-EGFR monoclonal antibody therapy for RAS wild-type metastatic colorectal cancer, Br. J. Cancer. 112 (2015) 1888–1894. doi:10.1038/BJC.2015.173.

[160] F. Caputo, C. Santini, C. Bardasi, K. Cerma, A. Casadei-Gardini, A. Spallanzani, K. Andrikou, S. Cascinu, F. Gelsomino, BRAF-Mutated Colorectal Cancer: Clinical and Molecular Insights, Int. J. Mol. Sci. 20 (2019). doi:10.3390/IJMS20215369.

[161] G. Cathomas, PIK3CA in Colorectal Cancer, Front. Oncol. 4 (2014). doi:10.3389/FONC.2014.00035.

[162] L. Huang, J. Carney, D.M. Cardona, C.M. Counter, Decreased tumorigenesis in mice with a Kras point mutation at C118, Nat. Commun. 2014 51. 5 (2014) 1–10. doi:10.1038/ncomms6410.

[163] P.S. W.De Roock , B.Claes , D.Bernasconi , J.De Schutter , B.Biesmans , G.Fountzilas , KT.Kalogeras , V.Kotoula, D.Papamichael , P. Laurent-Puig , F.Penault-Llorca , P.Rougier , B.Vincenzi , D.Santini , G.Tonini , F.Cappuzzo, M.Frattini , F.Molinari, Effects of KRAS, BRAF, NRAS, and PIK3CA mutations on the efficacy of cetuximab plus chemotherapy in chemotherapy-refractory metastatic colorectal cancer: a retrospective consortium analysis, Lancet. Oncol. 11 (2010) 753–762. doi:10.1016/S1470-2045(10)70130-3.

[164] R. STEC, A. SEMENIUK-WOJTAŚ, R. CHARKIEWICZ, L. BODNAR, J. KORNILUK, M. SMOTER, L. CHYCZEWSKI, J. NIKLIŃSKI, C. SZCZYLIK, Mutation of the PIK3CA gene as a prognostic factor in patients with colorectal cancer, Oncol. Lett. 10 (2015) 1423. doi:10.3892/OL.2015.3398.

[165] X. Liao, P. Lochhead, R. Nishihara, T. Morikawa, A. Kuchiba, M. Yamauchi, Y. Imamura, Z.R. Qian, Y. Baba, K. Shima, R. Sun, K. Nosho, J.A. Meyerhardt, E. Giovannucci, C.S. Fuchs, A.T. Chan, S. Ogino, Aspirin Use, Tumor PIK3CA Mutation, and Colorectal-Cancer Survival, Http://Dx.Doi.Org/10.1056/NEJMoa1207756. 367 (2012) 1596–1606. doi:10.1056/NEJMOA1207756.

[166] Y. Yan, Q. Guo, F. Wang, R. Adhikari, Z. Zhu, H. Zhang, W. Zhou, H. Yu, J. Li, J. Zhang, Cell-Free DNA: Hope and Potential Application in Cancer, Front. Cell Dev. Biol. 9 (2021) 639233. doi:10.3389/FCELL.2021.639233.

[167] Agostini M, Pucciarelli S, Enzo M, Del Bianco, Briarava M, Bedin C, Lonardi S, Maretto I, Friso M, Mescoli C, Circulating cell-free DNA: a promising marker of pathologic tumor response in rectal cancer patients receiving preoperative chemoradiotherapy, Ann. Surg. Oncol. 18 (2011) 2461–2468. doi:10.1245/S10434-011-1638-Y.

[168] Gormally E, Caboux E,Hainaut P. Vineis P, Circulating free DNA in plasma or serum

as biomarker of carcinogenesis: practical aspects and biological significance, Mutat. Res. 635 (2007) 105–117. doi:10.1016/J.MRREV.2006.11.002.

[169] A.F. Saraceni, R. Azevedo, C.M.G. Almeida, A.C. Baraviera, D.R. Kiss, M.G. Almeida, Association of fecal occult blood tests results with colonoscopic findings in a general hospital and validation of the screening test, J. Coloproctology. 39 (2019) 121–126. doi:10.1016/J.JCOL.2018.10.011.

[170] Wong MC, Ching JY, Ng SS Luk AK Chan VC, Lam TY, Sung JJ, Factors associated with false-positive and false-negative fecal immunochemical test results for colorectal cancer screening, Gastrointest. Endosc. 81 (2015) 596–607. doi:10.1016/J.GIE.2014.08.006.

[171] N. Singhal, M. Kumar, P.K. Kanaujia, J.S. Virdi, MALDI-TOF mass spectrometry: an emerging technology for microbial identification and diagnosis, Front. Microbiol. 6 (2015) 791.

[172] J.F. Riemann, C. Schröder, M. Kallenbach, K. Giersiepen, H.-J. Schmoll, Benefits and Risks of Colorectal Cancer Screening, Oncol. Res. Treat. 37 (2014) 11–20. doi:10.1159/000364886.

[173] M. Niederreiter, L. Niederreiter, A. Schmiderer, H. Tilg, A. Djanani, Colorectal cancer screening and prevention—pros and cons, Memo - Mag. Eur. Med. Oncol. 2019 123. 12 (2019) 239–243. doi:10.1007/S12254-019-00520-Z.

[174] G. Golshani, Y. Zhang, Advances in immunotherapy for colorectal cancer: a review:, Https://Doi.Org/10.1177/1756284820917527. 13 (2020). doi:10.1177/1756284820917527.

[175] C. Ben-Dov, B. Hartmann, J. Lundgren, J. Valcárcel, Genome-wide analysis of alternative pre-mRNA splicing, J. Biol. Chem. 283 (2008) 1229–1233. doi:10.1074/jbc.R700033200.

[176] Regulation of mRNA Splicing by Signal Transduction | Learn Science at Scitable, (n.d.). https://www.nature.com/scitable/topicpage/regulation-of-mrna-splicing-by-signal-transduction-14128469/ (accessed September 7, 2020).

[177] M.P. Paronetto, I. Passacantilli, C. Sette, Alternative splicing and cell survival: from tissue homeostasis to disease, Cell Death Differ. 2016 2312. 23 (2016) 1919–1929. doi:10.1038/cdd.2016.91.

[178] B. Roy, L.M. Haupt, L.R. Griffiths, Review: Alternative Splicing (AS) of Genes As An Approach for Generating Protein Complexity, Curr. Genomics. 14 (2013) 182. doi:10.2174/1389202911314030004.

[179] Y. Zhang, J. Qian, C. Gu, Y. Yang, Alternative splicing and cancer: a systematic review, Signal Transduct. Target. Ther. 2021 61. 6 (2021) 1–14. doi:10.1038/s41392-021-00486-7.

[180] W. Jiang, L. Chen, Alternative splicing: Human disease and quantitative analysis from high-throughput sequencing, Comput. Struct. Biotechnol. J. 19 (2021) 183. doi:10.1016/J.CSBJ.2020.12.009.

[181] L.M. do Canto, S.S. Cury, M.C. Barros-Filho, B.E.C. Kupper, M.D.F. de S. Begnami, C. Scapulatempo-Neto, R.F. Carvalho, F.A. Marchi, D.A. Olsen, J.S. Madsen, B.M. Havelund, S. Aguiar, S.R. Rogatto, Locally advanced rectal cancer transcriptomic-based secretome analysis reveals novel biomarkers useful to identify patients according to neoadjuvant chemoradiotherapy response, Sci. Reports 2019 91. 9 (2019) 1–11. doi:10.1038/s41598-019-45151-w.

[182] G. Pira, P. Uva, A.M. Scanu, P.C. Rocca, L. Murgia, E. Uleri, C. Piu, A. Porcu, C. Carru, A. Manca, I. Persico, M.R. Muroni, F. Sanges, C. Serra, A. Dolei, A. Angius, M.R. De Miglio, Landscape of transcriptome variations uncovering known and novel driver events in colorectal carcinoma, Sci. Reports 2020 101. 10 (2020) 1–12. doi:10.1038/s41598-019-57311-z.

[183] D. Black, Mechanisms of alternative pre-messenger RNA splicing, Annu. Rev. Biochem. 72 (2003) 291–336. doi:10.1146/annurev.biochem.72.121801.161720.

[184] M.C. Wahl, C.L. Will, R. Lührmann, The Spliceosome: Design Principles of a Dynamic RNP Machine, Cell. 136 (2009) 701–718. doi:10.1016/J.CELL.2009.02.009.

[185] R. Singh, RNA–Protein Interactions That Regulate Pre-mRNA Splicing, Gene Expr. 10 (2002) 79. /pmc/articles/PMC5977533/ (accessed July 12, 2021).

[186] RNA Splicing - an overview | ScienceDirect Topics, (n.d.). https://www.sciencedirect.com/topics/neuroscience/rna-splicing (accessed July 12, 2021).

[187] N. Toor, K.S. Keating, S.D. Taylor, A.M. Pyle, Crystal structure of a self-spliced group II intron, Science. 320 (2008) 77–82. doi:10.1126/SCIENCE.1153803.

[188] H. Dvinge, Regulation of alternative mRNA splicing: old players and new perspectives, FEBS Lett. 592 (2018) 2987–3006. doi:10.1002/1873-3468.13119.

[189] J. Han, J. Xiong, D. Wang, X.-D. Fu, Pre-mRNA splicing: where and when in the nucleus, Trends Cell Biol. 21 (2011) 336. doi:10.1016/J.TCB.2011.03.003.

[190] Y. Lee, D.C. Rio, Mechanisms and Regulation of Alternative Pre-mRNA Splicing, Annu. Rev. Biochem. 84 (2015) 291. doi:10.1146/ANNUREV-BIOCHEM-060614-

034316.

[191] La Cognata V, Iemmolo R, D'Agata V, Scuderi S, Drago F, Zappia M, Cavallaro S, Increasing the Coding Potential of Genomes Through Alternative Splicing: The Case of PARK2 Gene, (n.d.). https://www.researchgate.net/figure/The-alternative-splicing-mechanismThe-spliceosome-machinery-U1-U2-U4-U5-and-U6_fig1_263356754 (accessed July 28, 2021).

[192] M. Chen, J.L. Manley, Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches, 10 (2009). /pmc/articles/PMC2958924/ (accessed July 12, 2021).

[193] B.-G. Yue, Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine 926 Regulation of Adenovirus Alternative Pre-mRNA Splicing Functional Characterization of Exonic and Intronic Splicing Enhancer Elements, (2000).

[194] A. Larson, B.J. Fair, J.A. Pleiss, Interconnections Between RNA-Processing Pathways Revealed by a Sequencing-Based Genetic Screen for Pre-mRNA Splicing Mutants in Fission Yeast, G3 Genes|Genomes|Genetics. 6 (2016) 1513. doi:10.1534/G3.116.027508.

[195] Y. Zhou, Y. Lu, W. Tian, Epigenetic features are significantly associated with alternative splicing, BMC Genomics 2012 131. 13 (2012) 1–12. doi:10.1186/1471-2164-13-123.

[196] Z. Louadi, M. Oubounyt, H. Tayara, K.T. Chong, Deep Splicing Code: Classifying Alternative Splicing Events Using Deep Learning, Genes (Basel). 10 (2019). doi:10.3390/GENES10080587.

[197] L. Vigevani, A. Gohr, T. Webb, M. Irimia, J. Valcárcel, Molecular basis of differential 3′ splice site sensitivity to anti-tumor drugs targeting U2 snRNP, Nat. Commun. 2017 81. 8 (2017) 1–15. doi:10.1038/s41467-017-02007-z.

[198] G. Thill, V. Castelli, S. Pallud, M. Salanoubat, P. Wincker, P. de la Grange, D. Auboeuf, V. Schächter, J. Weissenbach, ASEtrap: A biological method for speeding up the exploration of spliceomes, Genome Res. 16 (2006) 776. doi:10.1101/GR.5063306.

[199] L.P. Iñiguez, M. Ramírez, W.B. Barbazuk, G. Hernández, Identification and analysis of alternative splicing events in Phaseolus vulgaris and Glycine max, BMC Genomics 2017 181. 18 (2017) 1–17. doi:10.1186/S12864-017-4054-2.

[200] A. Chaudhury, P. Chander, P.H. Howe, Heterogeneous nuclear ribonucleoproteins (hnRNPs) in cellular processes: Focus on hnRNP E1's multifunctional regulatory roles, RNA. 16 (2010) 1449. doi:10.1261/RNA.2254110.

[201] Y. WANG, J. LIU, B. HUANG, Y.-M. XU, J. LI, L.-F. HUANG, J. LIN, J. ZHANG, Q.-H. MIN, W.-M. YANG, X.-Z. WANG, Mechanism of alternative splicing and its regulation, Biomed. Reports. 3 (2015) 152. doi:10.3892/BR.2014.407.

[202] J. Ule, B.J. Blencowe, Alternative Splicing Regulatory Networks: Functions, Mechanisms, and Evolution, Mol. Cell. 76 (2019) 329–345. doi:10.1016/J.MOLCEL.2019.09.017.

[203] M.S. Wong, J.B. Kinney, A.R. Krainer, Quantitative activity profile and context dependence of all human 5′ splice sites, Mol. Cell. 71 (2018) 1012. doi:10.1016/J.MOLCEL.2018.07.033.

[204] S.S. Itskovich, A. Gurunathan, J. Clark, M. Burwinkel, M. Wunderlich, M.R. Berger, A. Kulkarni, K. Chetal, M. Venkatasubramanian, N. Salomonis, A.R. Kumar, L.H. Lee, MBNL1 regulates essential alternative RNA splicing patterns in MLL-rearranged leukemia, Nat. Commun. 2020 111. 11 (2020) 1–14. doi:10.1038/s41467-020-15733-8.

[205] P. Konieczny, E. Stepniak-Konieczna, K. Sobczak, MBNL proteins and their target RNAs, interaction and splicing regulation, Nucleic Acids Res. 42 (2014) 10873. doi:10.1093/NAR/GKU767.

[206] E. El Marabti, I. Younis, E. El Marabti, I. Younis, E. El Marabti, I. Younis, E. El Marabti, I. Younis, The Cancer Spliceome: Reprograming of Alternative Splicing in Cancer, Front. Mol. Biosci. 5 (2018) 80. doi:10.3389/FMOLB.2018.00080.

[207] D. Aghamirzaie, E. Collakova, S. Li, R. Grene, CoSpliceNet: a framework for co-splicing network inference from transcriptomics data, BMC Genomics 2016 171. 17 (2016) 1–16. doi:10.1186/S12864-016-3172-6.

[208] C. Haynes, L.M. Iakoucheva, Serine/arginine-rich splicing factors belong to a class of intrinsically disordered proteins, Nucleic Acids Res. 34 (2006) 305–312. doi:10.1093/NAR/GKJ424.

[209] A. Busch, K.J. Hertel, EVOLUTION OF SR PROTEIN AND HnRNP SPLICING REGULATORY FACTORS, Wiley Interdiscip. Rev. RNA. 3 (2012) 1. doi:10.1002/WRNA.100.

[210] S. Jeong, SR Proteins: Binders, Regulators, and Connectors of RNA, Mol. Cells. 40 (2017) 1. doi:10.14348/MOLCELLS.2017.2319.

[211] Y. Komeno, Y.-J. Huang, J. Qiu, L. Lin, Y. Xu, Y. Zhou, L. Chen, D.D. Monterroza, H. Li, R.C. DeKelver, M. Yan, X.-D. Fu, D.-E. Zhang, SRSF2 Is Essential for Hematopoiesis, and Its Myelodysplastic Syndrome-Related Mutations Dysregulate Alternative Pre-mRNA Splicing, Mol. Cell. Biol. 35 (2015) 3071.

doi:10.1128/MCB.00202-15.

[212] E. El Marabti, I. Younis, The Cancer Spliceome: Reprograming of Alternative Splicing in Cancer, Front. Mol. Biosci. 5 (2018) 80. doi:10.3389/FMOLB.2018.00080.

[213] S.A. Mabon, T. Misteli, Differential Recruitment of Pre-mRNA Splicing Factors to Alternatively Spliced Transcripts In Vivo, PLoS Biol. 3 (2005) 1893–1901. doi:10.1371/JOURNAL.PBIO.0030374.

[214] Han H, Irimia M, Ross PJ, Sung HK, Alipanahi B, David L, Golipour A, Gabut M, Michael IP, Nachman EN, Wang E, Trcka D, Thompson T, O'Hanlon D, Slobodeniuc V, Barbosa-Morais NL, Blencowe BJ,Burge CB, Moffat J, Frey BJ, Nagy A, Ellis J, MBNL proteins repress ES-cell-specific alternative splicing and reprogramming, Nature. 498 (2013) 241–245. doi:10.1038/nature12270.

[215] J. Choi, K.E. Personius, M. DiFranco, W. Dansithong, C. Yu, S. Srivastava, D.M. Dixon, D.B. Bhatt, L. Comai, J.L. Vergara, S. Reddy, Muscleblind-Like 1 and Muscleblind-Like 3 Depletion Synergistically Enhances Myotonia by Altering Clc-1 RNA Translation, EBioMedicine. 2 (2015) 1034–1047. doi:10.1016/J.EBIOM.2015.07.028.

[216] P.Y. Wang, K.T. Chang, Y.M. Lin, T.Y. Kuo, G.S. Wang, Ubiquitination of MBNL1 Is Required for Its Cytoplasmic Localization and Function in Promoting Neurite Outgrowth, Cell Rep. 22 (2018) 2294–2306. doi:10.1016/J.CELREP.2018.02.025.

[217] R. Roha, THE MECHANSIMS BY WHICH MBNL1 AND MBNL2 SEPARATE THEIR ALTERNATIVE-SPLICING FUNCTIONS, (2013).

[218] L.J. Sznajder, M. l. Michalak, K. Taylor, P. Cywoniuk, M. -l Kabza, A. Wojtkowiak-Szlachcic, M. Mat -loka, P. Konieczny, K. Sobczak, Mechanistic determinants of MBNL activity, Nucleic Acids Res. 44 (2016) 10326–10342. doi:10.1093/nar/gkw915.

[219] M. Pascual, M. Vicente, L. Monferrer, R. Artero, The Muscleblind family of proteins: an emerging class of regulators of developmentally programmed alternative splicing, Differentiation. 74 (2006) 65–80. doi:10.1111/J.1432-0436.2006.00060.X.

[220] F.Z. Xiong Y, Deng Y, Wang K, Zhou H, Zheng X, Si L, Profiles of alternative splicing in colorectal cancer and their clinical significance: A study based on large-scale sequencing data, EBioMedicine. 36 (2018) 183–195. doi:10.1016/J.EBIOM.2018.09.021.

[221] A. Bisognin, S. Pizzini, L. Perilli, G. Esposito, S. Mocellin, D. Nitti, P. Zanovello, S. Bortoluzzi, S. Mandruzzato, An integrative framework identifies alternative splicing events in colorectal cancer development, Mol. Oncol. 8 (2014) 129–141. doi:10.1016/J.MOLONC.2013.10.004.

[222] E. Stickeler, F. Kittrell, D. Medina, S.M. Berget, Stage-specific changes in SR splicing factors and alternative splicing in mammary tumorigenesis, Oncogene 1999 1824. 18 (1999) 3574–3582. doi:10.1038/sj.onc.1202671.

[223] L. S, C. C, Alternative RNA splicing and cancer., Wiley Interdiscip. Rev. RNA. 4 (2013) 547–566. doi:10.1002/WRNA.1178.

[224] J.H. Bahn, Alternative Splicing in Human Colorectal Cancer, (2010). https://trace.tennessee.edu/utk_graddiss (accessed July 12, 2021).

[225] A. Cerasuolo, L. Buonaguro, F.M. Buonaguro, M.L. Tornesello, The Role of RNA Splicing Factors in Cancer: Regulation of Viral and Human Gene Expression in Human Papillomavirus-Related Cervical Cancer, Front. Cell Dev. Biol. 8 (2020) 474. doi:10.3389/FCELL.2020.00474.

[226] M. Blanchette, R.E. Green, S.E. Brenner, D.C. Rio, Global analysis of positive and negative pre-mRNA splicing regulators in Drosophila, Genes Dev. 19 (2005) 1306–1314. doi:10.1101/GAD.1314205.

[227] J.D. Fackenthal, L.A. Godley, Aberrant RNA splicing and its functional consequences in cancer cells, Dis. Model. Mech. 1 (2008) 37. doi:10.1242/DMM.000331.

[228] W. Li, X. Peng, J. Lang, C. Xu, Targeting Mouse Double Minute 2: Current Concepts in DNA Damage Repair and Therapeutic Approaches in Cancer, Front. Pharmacol. 11 (2020). doi:10.3389/FPHAR.2020.00631.

[229] Li D, Xie K, Zhang L, Yao X, Li H, Xu Q, Wang X, Jiang J, Dual blockade of vascular endothelial growth factor (VEGF) and basic fibroblast growth factor (FGF-2) exhibits potent anti-angiogenic effects, Cancer Lett. 377 (2016) 164–173. doi:10.1016/J.CANLET.2016.04.036.

[230] M. De Rosa, G. Morelli, E. Cesaro, F. Duraturo, M. Turano, G.B. Rossi, P. Delrio, P. Izzo, Alternative splicing and nonsense-mediated mRNA decay in the regulation of a new adenomatous polyposis coli transcript, Gene. 395 (2007) 8–14. doi:10.1016/J.GENE.2006.10.027.

[231] Luo.F, Ye.H, Hamoudi.R, Dong.G, Zhang.W, Patek.CE, Poulogiannis.G, Arends.MJ, K-ras exon 4A has a tumour suppressor effect on carcinogen-induced murine colonic adenoma formation, J. Pathol. 220 (2010) 542–550. doi:10.1002/PATH.2672.

[232] H. Zhang, R.L. Brown, Y. Wei, P. Zhao, S. Liu, X. Liu, Y. Deng, X. Hu, J. Zhang, X.D. Gao, Y. Kang, A.M. Mercurio, H.L. Goel, C. Cheng, CD44 splice isoform switching determines breast cancer stem cell state, Genes Dev. 33 (2019) 166. doi:10.1101/GAD.319889.118.

[233] R. Karni, E. de Stanchina, S.W. Lowe, R. Sinha, D. Mu, A.R. Krainer, The gene encoding the splicing factor SF2/ASF is a proto-oncogene, Nat. Struct. Mol. Biol. 14 (2007) 185. doi:10.1038/NSMB1209.

[234] L.M. Urbanski, N. Leclair, O. Anczuków, Alternative-splicing defects in cancer: Splicing regulators and their downstream targets, guiding the way to novel cancer therapeuticsAlternative Splicing, (2018). doi:10.1002/wrna.1476.

[235] D. Pradella, C. Naro, C. Sette, C. Ghigna, EMT and stemness: flexible processes tuned by alternative splicing in development and cancer progression, Mol. Cancer. 16 (2017) 8. doi:10.1186/s12943-016-0579-2.

[236] I.M. Shapiro, A.W. Cheng, N.C. Flytzanis, M. Balsamo, J.S. Condeelis, M.H. Oktay, C.B. Burge, F.B. Gertler, An EMT–Driven Alternative Splicing Program Occurs in Human Breast Cancer and Modulates Cellular Phenotype, PLoS Genet. 7 (2011) e1002218. doi:10.1371/journal.pgen.1002218.

[237] A. Sánchez, I. Fabregat, T. Mitaka, Growth factor-and cytokine-driven pathways governing liver stemness and differentiation, (2010). doi:10.3748/wjg.v16.i41.5148.

[238] I. Fabregat, D. Caballero-Díaz, Transforming Growth Factor-β-Induced Cell Plasticity in Liver Fibrosis and Hepatocarcinogenesis, Front. Oncol. 8 (2018) 357. doi:10.3389/FONC.2018.00357.

[239] H. Zheng, Y. Kang, Multilayer control of the EMT master regulators, Oncogene. 33 (2014) 1755–1763. doi:10.1038/onc.2013.128.

[240] A.G. De Herreros, S. Peiró, M. Nassour, P. Savagner, Snail family regulation and epithelial mesenchymal transitions in breast cancer progression, J. Mammary Gland Biol. Neoplasia. 15 (2010) 135. doi:10.1007/S10911-010-9179-8.

[241] I.Muqbila, J.Wua, S.Azmi, A.Aboukameela, R.M.Mohammadab, A. S.Azmi, Snail nuclear transport: The gateways regulating epithelial-to-mesenchymal transition?, (n.d.). https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/transcription-factor-snail (accessed July 12, 2021).

[242] S. Vega, A. V. Morales, O.H. Ocaña, F. Valdés, I. Fabregat, M.A. Nieto, Snail blocks the cell cycle and confers resistance to cell death, Genes Dev. 18 (2004) 1131–1143. doi:10.1101/GAD.294104.

[243] RL.Brown, LM.Reinke, MS.Damerow, D.Perez, LA.Chodosh, J.Yang, C.Cheng, CD44 splice isoform switching in human and mouse epithelium is essential for epithelial-mesenchymal transition and breast cancer progression, J. Clin. Invest. 121 (2011) 1064–1074. doi:10.1172/JCI44540.

[244] G. Biamonti, S. Bonomi, S. Gallo, C. Ghigna, Making alternative splicing decisions during epithelial-to-mesenchymal transition (EMT), Cell. Mol. Life Sci. 2012 6915. 69 (2012) 2515–2526. doi:10.1007/S00018-012-0931-7.

[245] A.B. James, N.H. Syed, S. Bordage, J. Marshall, G.A. Nimmo, G.I. Jenkins, P. Herzyk, J.W.S. Brown, H.G. Nimmo, Alternative Splicing Mediates Responses of the Arabidopsis Circadian Clock to Temperature Changes, Plant Cell. 24 (2012) 961. doi:10.1105/TPC.111.093948.

[246] N. Navvabi, P.Kolikova, P. Hosek, F. Zitricky, A.Navvabi, O. Vycital, J.Bruha, R. Palek, J. Rosendorf, V. Liska, P. Pitule, Altered Expression of MBNL Family of Alternative Splicing Factors in Colorectal Cancer, Cancer Genomics Proteomics. 18 (2021) 295. doi:10.21873/CGP.20260.

[247] K. Charizanis, K.-Y. Lee, R. Batra, M. Goodwin, C. Zhang, Y. Yuan, L. Shiue, M. Cline, M.M. Scotti, G. Xia, A. Kumar, T. Ashizawa, H.B. Clark, T. Kimura, M.P. Takahashi, H. Fujimura, K. Jinnai, H. Yoshikawa, M. Gomes-Pereira, G. Gourdon, N. Sakai, S. Nishino, T.C. Foster, M. Ares, Jr, R.B. Darnell, M.S. Swanson, Muscleblind-Like 2 Mediated Alternative Splicing in the Developing Brain and Dysregulation in Myotonic Dystrophy, Neuron. 75 (2012) 437. doi:10.1016/J.NEURON.2012.05.029.

[248] M.Gabut, P.Samavarchi-Tehrani, Wang X, V.Slobodeniuc, D. O'Hanlon, HK. Sung, M. Alvarez, S. Talukder, Q. Pan, EO. Mazzoni, S. Nedelec, H. Wichterle, K. Woltjen, TR. Hughes, PW. Zandstra, A. Nagy, JL. Wrana, BJ. Blencowe, An alternative splicing switch regulates embryonic stem cell pluripotency and reprogramming, Cell. 147 (2011) 132–146. doi:10.1016/J.CELL.2011.08.023.

[249] Y. Kino, C. Washizu, Y. Oma, H. Onishi, Y. Nezu, N. Sasagawa, N. Nukina, S.Ishiura, MBNL and CELF proteins regulate alternative splicing of the skeletal muscle chloride channel CLCN1, Nucleic Acids Res. 37 (2009) 6477–6490. doi:10.1093/NAR/GKP681.

[250] C. Chen, S. Zhao, A. Karnad, J.W. Freeman, The biology and role of CD44 in cancer progression: therapeutic implications, J. Hematol. Oncol. 2018 111. 11 (2018) 1–23. doi:10.1186/S13045-018-0605-5.

[251] B. Bánky, L. Rásó-Barnett, T. Barbai, J. Tímár, P. Becságh, E. Rásó, Characteristics of CD44 alternative splice pattern in the course of human colorectal adenocarcinoma progression, Mol. Cancer. 11 (2012) 83. doi:10.1186/1476-4598-11-83.

[252] F. Holm, E. Hellqvist, C.N. Mason, S.A. Ali, N. Delos-Santos, C.L. Barrett, H.-J. Chun, M.D. Minden, R.A. Moore, M.A. Marra, V. Runza, K.A. Frazer, A. Sadarangani, C.H.M. Jamieson, Reversion to an embryonic alternative splicing program enhances leukemia stem cell self-renewal, (2021). doi:10.1073/pnas.1506943112.

[253] M. Todaro, M. Gaggianesi, V. Catalano, A. Benfante, F. Iovino, M. Biffoni, T. Apuzzo, I. Sperduti, S. Volpe, G. Cocorullo, G. Gulotta, F. Dieli, R. De Maria, G. Stassi, CD44v6 is a marker of constitutive and reprogrammed cancer stem cells driving colon cancer metastasis, Cell Stem Cell. 14 (2014) 342–356. doi:10.1016/j.stem.2014.01.009.

[254] S.J. Wang, G. Wong, A.-M. de Heer, W. Xia, L.Y. Bourguignon, CD44 variant isoforms in head and neck squamous cell carcinoma progression, Laryngoscope. 119 (2009) 1518. doi:10.1002/LARY.20506.

[255] D.B. Bernkopf, E.D. Williams, Potential role of EPB41L3 (Protein 4.1B/Dal-1) as a target for treatment of advanced prostate cancer, Expert Opin. Ther. Targets. 12 (2008) 845–853. doi:10.1517/14728222.12.7.845.

[256] R. Zeng, Y. Liu, ZJ. Jiang, JP. Huang, Y. Wang, XF.Li, WB. Xiong, XC. Wu, JR. Zhang, QE. Wang, YF. Zheng, EPB41L3 is a potential tumor suppressor gene and prognostic indicator in esophageal squamous cell carcinoma, Int. J. Oncol. 52 (2018) 1443–1454. doi:10.3892/IJO.2018.4316.

9   Acknowledgements

10   Publications

My main research article "Altered Expression of MBNL Family of Alternative Splicing Factors in Colorectal Cancer" related to the topic of my study was published in Cancer Genomics Proteomics Journal in May-Jun 2021 and it forms a core of my submitted thesis. During the work on the project I had a chance to get familiar with multiple methods like DNA and RNA isolation protocols, PCR or qPCR analysis. All the knowledge I acquired I applied during my regular summer stay collaborations in my home country, Iran.

Additional two publications as coauthor originate from my cooperation with Pavel Vodicka, MD, PhD, DSc, Institute of Experimental Medicine. Molecular biology of cancer group in Prague. "Oxidative Damage in Sporadic Colorectal Cancer: Molecular Mapping of Base Excision Repair Glycosylases MUTYH and hOGG1 in Colorectal Cancer Patients" and "MiR-140 leads to MRE11 downregulation and ameliorates oxaliplatin treatment and therapy response in CRC patients".

Additionally, I have other papers from my cooperation with other colleagues and although they are not thematically directly connected to my thesis topic, during their preparation I used current and gained additional experience – in particular, first publication used diagnostic PCR tests for detection of HBsAg confirmed HBV infection and second focused on marine celluloses and their biotechnological significance from industrial perspectives, that can serve as basis for greater efficacy and specificity for the therapeutics, as it is estimated that more than 60% of commercially available anticancer drugs are natural biomimetic inspired and among the marine organisms, algae were revealed to be one of the major sources of new compounds of marine origin, including those exhibiting antitumor and cytotoxic potential. These compounds demonstrated ability to mediate specific inhibitory activities on a number of key cellular processes, including apoptosis pathways, angiogenesis, migration and invasion, in

both *in vitro* and in vivo models. As mentioned above, for these publications I used my knowledge acquired at the Faculty of Medicine in Pilsen, mainly in PCR and qPCR, I was expanded my skillset by other methods like ELISA and immunochromatography. In addition to published work, I dedicated significant amount of time to additional projects related to colorectal cancer, namely role of TRK protein family in CRC progression and effect of EMT induction on alternative splicing. Unfortunately, but in line with a spirit of research, preliminary data in both topics showed no relation of analyzed genes to expected effects and therefore after agreement with my supervisor we decided not to proceed in those topics and focused on main project described in presented thesis.

List of publications

Nazila Navvabi *, Pavla Kolikova, Petr Hosek, Frantisek Zitricky, Azita Navvabi, Ondrej Vycital, Jan Bruha, Richard Palek, Jachym Rosendorf, Vaclav Liska, Pavel Pitule

Cancer Genomics Proteomics. May-Jun 2021;18(3):295-306. doi: 10.21873/cgp.20260. **"Altered Expression of MBNL Family of Alternative Splicing Factors in Colorectal Cancer"**

*Contributed as the first author and was involved in study design, performed experiments, analysis and writing of the paper

Miriam J. Kavec, Marketa Urbanova, Pavol Makovicky, Alena Opattova, Kristyna Tomasova, Michal Kroupa, Klara Kostovcikova, Anna Siskova, **Nazila Navvabi***, Michaela Schneiderova,Veronika Vymetalkova, Ludmila Vodickova and Pavel Vodicka

International Journal of Molecular Sciences, 2022 May 20;23(10):5704 doi:10.3390/ijms 23105704.

 "Oxidative Damage in Sporadic Colorectal Cancer: Molecular Mapping of Base Excision Repair Glycosylases MUTYH and hOGG1 in Colorectal Cancer Patients"

*Contributed as the coauthor, participating on the experiments and analyses.

Josef Horak, Alexandra Dolnikova, Ozge Cuma, Andrea Cumova, **Nazila Navvabi** *, Ladislav Andera, Miroslav Levy, Vaclav Liska, Pavel Vodicka, and Alena Opattova,

Frontiers in Oncology, 022 Oct 17;12:959407. doi: 10.3389/fonc.2022.959407.2022

"MiR-140 leads to MRE11 downregulation and ameliorates oxaliplatin treatment and therapy response in CRC patients"

*Contributed as the coauthor.

Jiri Hatina, **Nazila Navvabi** *, Michaela Kripnerova, Katerina Houfkova, Martina Hajduskova, Natalie Havlíckova, and Martin Pesta

Springer Nature, chapter December 2022

"Recent Progress in Urothelial Bladderb Carcinoma: Basic Biology, Molecular Characterization, Conventional and Innovative Therapies and Tumour–Stroma Interactions"

*Contributed as the coauthor.

Articles related to the topic only vaguely or due to the methodologic impact

Nazila Navvabi*, Mohamad.H Khadem Ansarib, Azita Navvabi, H.R. Chalipa, Frantisek Zitricky

The Revista de Gastroenterología. November 2021; 10.1016/j.rgmx.2020.12.003

"Comparative assessment of Immunochromatography and Elisa diagnostic tests for detection of HBsAg in PCR confirmed HBV infection"

*Contributed as the first author and was involved in study design, performed experiments, analysis and writing and editation of the paper

Azita Navvabi, Brett I. Pletschke, **Nazila Navvabi***, Ahmad Homaei

Current Pharmaceutical Design

PMID: 35388747 .DOI: 10.2174/1381612828666220406125132 April 2022;

"**Marine celluloses and their biotechnological significance from industrial perspectives**"

*Contributed as the co- author and was involved in study design, writing and editing of the paper