VNIVERSITAT DE VALÈNCIA

Facultat de Medicina i Odontologia
Programa 3139 Departament de Medicina

A multi-omic liquid biopsy-based signature as a valuable tool to assess minimal residual disease in localised colorectal cancer

PhD Thesis presented by:
Noelia Tarazona Llavero

Thesis Director:
Prof. Andrés Cervantes Ruipérez

May, 2019

This thesis has been supported by the founding received by Noelia Tarazona - a grant by the European Society of Medical Oncology and a Río Hortega Contract (CM15/00246) from Instituto de Salud Carlos III- and also by PI15/02180 and PI18/01909 grants from Fondo de Investigaciones Sanitarias (Instituto de Salud Carlos III) from the Spanish Government.

Science is organized knowledge. Wisdom is organized life
Immanuel Kant

Agradecimientos

Hacen falta años para percatarse de que el no ser desagradecido es ya lograr bastante felicidad en este mundo. Miguel Delibes

…Desde hace años me pareció siempre la parte más importante y emotiva de la Tesis, y es curioso que ahora uno no sepa por dónde empezar, pero sencillamente empezaré por el principio. Esta Tesis no hubiera sido posible sin ti, hace 11 años que nos conocimos, por aquel entonces empezaba la residencia de Oncología una chica con inquietudes y con miedos pero que poco a poco y gracias a la gente que forma parte de este maravilloso equipo se dio cuenta de que la Oncología era mucho más que una especialidad, era una forma de vida. Me enseñaste en la consulta a plantearme preguntas, a ser crítica y a ser no sólo buen oncólogo sino mejor persona con los pacientes que al final es lo gratificante de esta especialidad, ellos. Gracias Andrés Cervantes, por tu mentorship, por tu confianza, por creer en mi en los buenos momentos y también en mis malas decisiones, pero sobre todo gracias por transmitirme que, y cito textualmente, ¨la medicina te tiene que apasionar¨. ¨Las personas inteligentes tienen un derecho sobre las ignorantes: el derecho de instruirlas¨ Ralph Waldo Emerson.

Juanmi Cejalvo, gracias por estar ahí, volvemos a coincidir tras años separados, pues decidimos que hacer investigación era lo que nos gustaba. No dudamos en buscar las respuestas alejados de nuestra zona de confort, pero aquí estamos de nuevo como al principio, más mayores, pero con las mismas ganas e ilusiones y dedicándonos a lo que nos apasiona. Gracias amigo, por escucharme, por haber estado en tantos momentos de mi vida que no olvidaré y por seguir estando aquí. Esta Tesis también es tuya porque fueron muchas las veces que fantaseamos con este momento, y al final todo llega. ¨Ese que habla tanto está completamente hueco, ya sabes que el cántaro vacío es el que más suena¨.

Y tú, que siempre estás a mi lado sin hacer ruido, Valentina Gambardella, probablemente sin tu apoyo incondicional, sin tus consejos, sin tu ayuda a nivel

laboral y personal no estaría escribiendo estas palabras como broche final a mis años de doctorado. Eres la mejor compañera de vida que uno puede tener. Ojalá te vieras como te vemos el resto, porque sencillamente eres brillante. Estés donde estés, que espero que sea a mi lado, serás parte de lo que soy. ¨Usted no sabe cómo yo valoro su sencillo coraje de quererme¨ Mario Benedetti.

Marisol Huerta, gracias por tu esfuerzo cada semana en este proyecto. Sigilosa, metódica y sencilla, es lo que te hace diferente y especial. Me encanta llegar al hospital cada mañana y saber que el primer rostro que veré será el tuyo. Gracias por ser como eres, por tu generosidad y por ser parte de la familia.

A mis amigas Laura, Marta y Carmen porque, aunque a veces me ausento seguís ahí. Gracias por admirarme, yo también os admiro a vosotras.

A mis amigos del cole, el otro día me recordabais que en 2º de EGB ya quería ser médico, gracias por seguir creciendo juntos.

A Ana Lluch, por tu ejemplo de trabajo y de trato humano. Aprendimos que, con naturalidad, carisma, trabajo constante y sin perderse uno en las tentaciones acabas convirtiéndote en la persona que uno ha soñado desde pequeño. Siempre serás recordada por la mujer que revolucionó el cáncer de mama en España y para los que te hemos tenido cerca por una mujer que se hizo así misma y transmitió su amor por la Oncología.

A Paco, por tener siempre una sonrisa, por tirar de mi cuando lo he necesitado, por enseñarme que la vida es sencilla si te rodeas de las personas que merecen la pena y, por esos fines de semana juntos en el laboratorio, comiendo comida no healthy con nuestro Netflix pero haciendo lo que tanto nos gusta, ciencia. Te deseo una vida plena y llena de éxitos.

A Pepa Castillo, por tu ejemplo de constancia, tu ayuda y tus consejos, y por hacerlo fácil siempre.

A mis compañeros del hospital Begoña, Tania, Pablo, Maite, Isabel, Desam, Amelia, Susana, Paloma, Gema, Cris y tantos otros, gracias por vuestro apoyo y por hacer que venir a trabajar merezca la pena.

A aquellos a los que no he mencionado, gracias, también sois participes del punto en el que estoy.

Finalmente, gracias a mi, por perseverar y a ti, si alguna vez lo lees, por hacerme fuerte y por crearme, sólo le diré una cosa: ¨Discúlpeme, no le había reconocido: he cambiado mucho¨.

Y a mis padres.

Y ahora sí, el comienzo de mi Tesis, espero que les guste.

Index

2. **Material and methods**

2.1 Study design

2.2 Patient population

2.3 Pathology

2.4 Custom gene panel design

2.5 Processing and DNA/RNA extraction from tumor tissue

2.6 Processing of plasma and extraction of circulating DNA

2.7 Processing and RNA extraction from fresh frozen tissue

2.8 DNA and RNA quantification from tissue and/or plasma

2.9 Targeted NGS

2.10 Data analysis and classification of variants

2.11 ctDNA analysis

2.12 ELISA (Enzyme-Linked Immunosorbent Assay): quantitative measurement of IL-6, IL-11 and TGFb in plasma.

2.13 NanoString nCounter gene expression assays-modified protocol

2.14 Statistical analysis

3. **Results and discussion**

3.1 Design of a targeted gene panel for ctDNA by bioinformatic screening of gene sets published in CRC. (**Aim 1**)

3.1.1 Assessment of the custom NGS genomic panel to detect mutations in patients with localized CRC.

3.2 Assessment and detection of MRD using circulating tumor DNA by droplet digital PCR in localized CRC patients and its correlation with clinical-pathological parameters. (**Aim 2**)

3.3 Evaluation of the prognostic and predictive value of a plasma proteomic panel. (**Aim 3**)

3.3.1 Determination of cytokine levels involved in the development of metastases (IL6, IL-11, and TGF-beta) in plasma samples.

3.3.2 Evaluation of the association between plasma levels of secreted factors with prognosis.

3.4 Classification of patients according to molecular subtypes (CRCAssigner and Consensus Molecular Subtypes (CMS) using a customized 38-gene assay for nCounter platform (NanoString Technology) (**Aim 4**)

3.4.1 Correlation of the molecular subtypes with ctDNA in patients with localized CRC.

3.5 Integration of a comprehensive molecular assessment to elaborate a valuable prognostic/predictive multi-omic model. (**Aim 5**)

4. **Conclusions**

5. **Summary**

6. **References**

Abbreviations

The following abbreviations are used in this manuscript:

CRC colorectal cancer
HDI human development index
DNA deoxyribonucleic acid
MSI microsatellite instability
PCR polymerase chain reaction
APC adenomatous polyposis coli
MUTYH mutY DNA glycosylase
IBD inflammatory bowel disease
NSAID nonsteroidal anti-inflammatory drug
CIMP CpG island methylator phenotype
TGFBR2 transforming growth factor-β receptor 2
PIK3CA phosphatidylinositol -4,5-biphospate 3-kinase catalytic subunit-α
ARID1A AT-rich interactive domain 1A
SOX9 SRY (sex-determining region Y) box 9 (SOX9)
FAM123B family with sequence similarity 123B
EGF epidermal growth factor
MAPK mitogen activated protein kinase
PI3K phosphatidylinositol 3-kinase
SEPT9 septin 9
TME tumor microenvironment
CT computed tomography
CEA carcinoembryonic antigen
TCGA The Cancer Genome Atlas
RNA ribonucleic acid
MMR mismatch repair
POLE polymerase ε
CMS consensus molecular subtype
MSS microsatellite stable

SCNA somatic copy number alteration
CIN chromosomal instability
EMT epithelial-mesenchymal transition
NCI National Cancer Institute
5-FU 5- fluorouracil
FDA Food and Drug Administration
mCRC metastatic colorectal cancer
SNP single-nucleotide polymorphism
HPGD hydroxyprostaglandin dehydrogenase
CDX2 caudal type homeobox transcription factor 2
DFS disease-free survival
HER2 Human epidermal growth factor receptor 2
cfDNA circulating free nucleic acids
ctDNA circulating tumor DNA
DFS disease-free survival
MRD minimal residual disease
CAFs Cancer associated Fibroblasts
TAMs tumor-associated macrophages
FFPE formalin-fixed and paraffin-embedded
NGS next generation sequencing
QC quality control
VAF variant allele frequency

INTRODUCTION

1.1 Background

Colorectal cancer (CRC) is a major public health problem, representing the third most frequently diagnosed cancer worldwide and the fourth leading cause of cancer death in the world. In 2012, the World Health Organization GLOBOCAN database estimated almost 1,4 million new cases of CRC across the world and 447,136 new cases of CRC in Europe (represented the 2nd most common cancer). The estimated prevalence of CRC was 3,543,582 and 1,203,943 worldwide and in Europe at 5 years, respectively [1,2].

CRC is a disease of modernity and the highest rates of incidence are detectable in developed countries. Moreover, as more people shift to Western diets and lifestyles, the incidence of CRC is likely to increase also in developing countries [3,4].

Rapid increases in both CRC incidence and mortality are now observed in many medium-to-high human development index (HDI) countries particularly in Eastern Europe, Asia and South America [3]. In contrast, CRC incidence and mortality rates are stable or declining in a number of the highest indexed HDI countries: the USA, Australia, New Zealand and several Western European countries. The reasons for the recent declining trends in incidence in these countries are imprecise and likely numerous but may partially reflect increased early detection and prevention through polypectomy (at least in the USA) obtained thanks to efficient screening programs. Together with the factors that have brought about declines in incidence, improvements in perioperative care, as well as chemotherapy and radiotherapy, will have contributed to the uniformly decreasing trends in CRC mortality in many high-income settings [5,6].

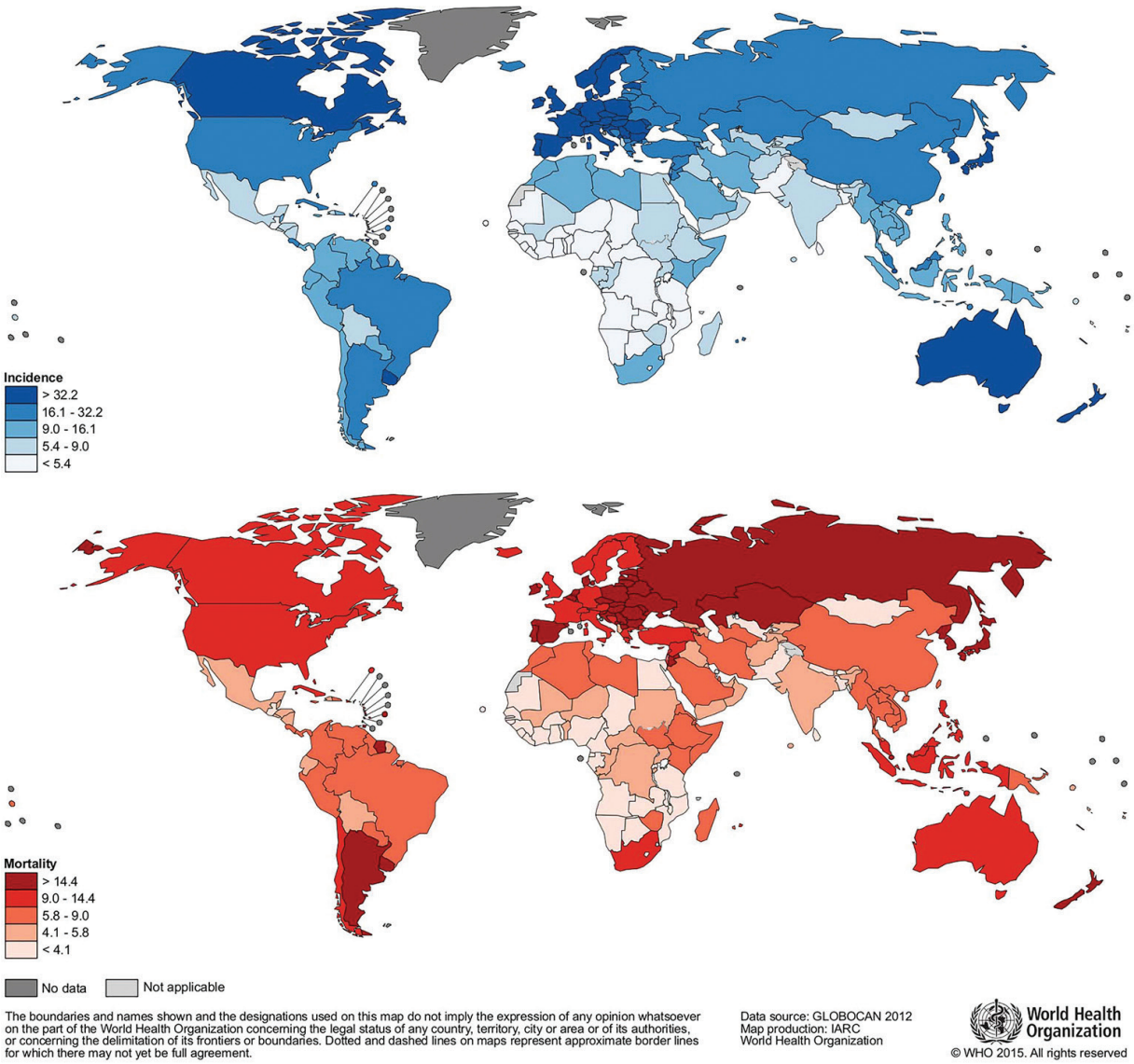


Figure 1. Worldwide colorectal cancer incidence and mortality rates (age adjusted according to the world standard population, per 100 000) in males in 2012 (GLOBOCAN 2012[2]).

Given the temporal profiles and demographic projections, the global burden of CRC is expected to increase by 60% to more than 2.2 million new cases and 1.1 million cancer deaths by 2030 [2].

Risk factors

Both genetic and environmental factors play an important part in the aetiology of CRC. The majority of CRC are sporadic; approximately three-quarters of patients have a negative family history. In most western populations, the average lifetime risk

for CRC is in the range of 3-5%. However, this risk almost doubles in individuals with a first-degree family member with CRC who was diagnosed at 50-70 years of age; the risk triples if the first-degree relative was <50 years of age at diagnosis. Risk further increases in individuals who have two or more affected family members. For sporadic CRC, this increased risk in the presence of affected family at least in part reflects low-penetrance genetic factors. Accordingly, positive family history has a role in approximately 15-20% of patients with CRC.

Indeed, a specific subgroup of the patient population is formed by those affected by a hereditary CRC syndrome, accounting for 5-10% of all patients. The most common syndrome in this category is Lynch syndrome. This syndrome is caused by a mutation in one of the DNA mismatch-repair genes: *MLH1*, *MSH2*, *MSH6*, *PMS2* or *EPCAM*. Impaired mismatch repair during replication gives rise to the accumulation of DNA mutations, which occur, in particular, in microsatellite DNA fragments with repetitive nucleotide sequences. This microsatellite instability (MSI) can be identified by polymerase chain reaction (PCR) testing. Patients with Lynch syndrome used to be identified by clinicopathological criteria, such as the Amsterdam and Bethesda criteria [7,8]. However, due to the possible use of immunotherapy in this setting, clinical practice is shifting towards unrestricted testing of tumor material of all patients diagnosed before 70 years of age by MSI PCR and immunohistochemistry for lack of expression of specific mismatch-repair proteins [9].

The second-most common hereditary CRC syndrome is familial adenomatous polyposis. This syndrome is caused by mutations in the adenomatous polyposis coli (*APC*) gene, which controls the activity of the WNT signaling pathway [7]. Most patients with familial adenomatous polyposis develop very large numbers of colorectal adenomas and subsequent CRC at a young age. Other hereditary CRC syndromes include polyposis associated with mutations in the mutY DNA glycosylase (*MUTYH*) gene, Peutz-Jeghers syndrome, serrated polyposis and juvenile polyposis.

Chronic colitis due to inflammatory bowel disease (IBD) is also associated with increased risk of CRC [10]. IBD explains only 1% of CRC in western populations, and a range of studies suggest that the incidence of CRC in those with IBD is decreasing because of effective anti-inflammatory treatments and improved surveillance, although this observation is not yet universal [11].

A range of environmental lifestyle factors influence the risk of developing CRC. The risk is increased by smoking, alcohol intake and increased body weight. With each unit increase of the body mass index, the risk for CRC increases by 2-3% [12]. Moreover, patients with type 2 diabetes mellitus also have an increased risk of CRC [13]. Moderate alcohol consumption (2–3 units per day) has been estimated to increase risk by 20%, whereas even higher alcohol consumption is associated with an up to 50% increased risk [14]. Prolonged heavy smoking has an effect of similar magnitude [15]. Intake of red meat and processed meat increases the risk of CRC by an estimated 1.16-fold per 100 g increase of daily intake [16]. By contrast, the consumption of milk, whole grains, fresh fruits and vegetables, as well as an intake of calcium, fibre, multivitamins and vitamin D, decrease the risk of CRC. Daily physical activity for 30 minutes has a similar magnitude of effect [17]. Low-dose aspirin has also been associated with decreased risk of CRC [18].

The prevalence of these modifiable lifestyle factors can explain, to a considerable extent, the geographical and socioeconomic differences in CRC incidence [19]. Several studies have estimated that 16-71% of CRCs in Europe and the United States are attributable to lifestyle factors [20]. Any benefit from lifestyle changes can be augmented by regular intake of aspirin and other NSAIDs [18]; however, this effect seems to depend on the host genotype [21]. Statin use might have a small preventive effect on the incidence of CRC [22], as does hormone therapy in postmenopausal women [23].

The various environmental factors that influence colorectal carcinogenesis is probably reflected in the heterogeneity of CRC, and has stimulated research into the

field of ‘molecular pathological epidemiology’, which focuses on the correlation between environmental and genetic factors, and between molecular tumor characteristics and disease progression [24]. Further research into the correlation between colonic microbiota and CRC will probably provide further insights.

Mechanisms/pathophysiology

The environmental and genetic factors that cause CRC are able to produce it by promoting the acquisition of hallmark behaviors of cancer in colon epithelial cells [25]. On way these hallmark cancer traits are acquired is through the progressive accumulation of genetic mutations and epigenetic alterations that activate oncogenes and inactivate tumor suppressor genes. The loss of genomic and/or epigenomic stability has been observed in the majority of early neoplastic lesions in the colon (namely, aberrant crypt foci, adenomas and serrated polyps) and is probably a central molecular and pathophysiological event in the initiation and formation of CRC [26,27]. The loss of genomic and epigenomic stability accelerates the accumulation of mutations and epigenetic alterations in oncogenes and tumor suppressor genes, which drive the malignant transformation of colon cells through rounds of clonal expansion that select for those cells with the most aggressive and malignant behavior [28-30]. A prevailing paradigm is that the cell of origin of most CRCs is a stem cell or stem cell-like cell that resides in the base of the colon crypts [31]. In this model, mutations in oncogenes and tumor suppressor genes in these cells lead to the formation of cancer stem cells, which are essential for the initiation and maintenance of a tumor.

In the colon, the evolution of normal epithelial cells to adenocarcinoma by and large follows a predictable progression of histological and concurrent epigenetic and genetic changes (Figure 2). In the ‘classic’ CRC formation model, the vast majority of cancers arise from a polyp beginning with an aberrant crypt, which then evolves into an early adenoma (<1 cm in size, with tubular or tubulovillous histology). The adenoma then progresses to an advanced adenoma (>1 cm in size, and/or with

villous histology) before finally becoming a CRC. This process is driven by the accumulation of mutations and epigenetic alterations and takes 10-15 years to occur but can progress more rapidly in certain settings (for example, in patients with Lynch syndrome) [32]. Notably, although the histology of conventional tubular adenomas is fairly homogeneous, the molecular biology of these polyps is heterogeneous, which might explain why some adenomas progress to CRC (approximately 10% of polyps) and some do not [33,34].

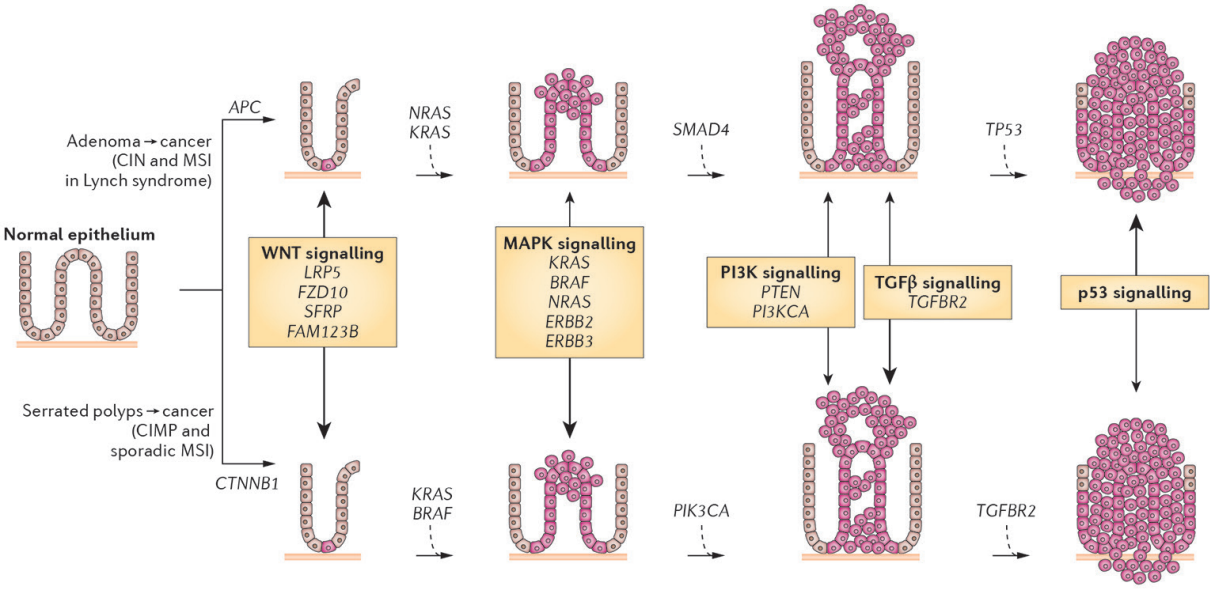


Figure 2. The polyp to colorectal cancer sequences.

Until 5-10 years ago, tubular and tubulovillous adenomatous polyps were thought to be the only lesions that were capable of progressing to cancer. However, some CRCs have been shown to evolve from a subset of polyps called sessile serrated polyps, which account for approximately 5-10% of all polyps. These serrated polyps arise by molecular and histological events that are distinct from tubular adenomas [35] and are classified into three categories: hyperplastic polyps, sessile serrated adenomas and traditional serrated adenomas [36]. The sessile serrated polyps have the potential to transform into CRCs through the following sequence: hyperplastic

polyp to sessile serrated polyp to adenocarcinoma. Furthermore, serrated polyps that arise in the right colon (which includes the caecum, ascending colon and transverse colon) commonly show MSI and a form of epigenetic instability characterized by excessive aberrant CpG island DNA methylation, termed the CpG island methylator phenotype (CIMP). By contrast, polyps that arise in the left colon (which includes the descending colon, sigmoid colon and rectum) are typically microsatellite stable but frequently carry mutations in KRAS, and a subset of these polyps have an attenuated form of the CIMP [37,38].

Given these molecular differences in the polyps and cancers they evolve into, a classification system for CRC has been proposed, with four subgroups of differing molecular features: *hypermutable microsatellite unstable, hypermutable microsatellite stable, microsatellite stable or chromosome unstable, and CIMP cancers*. The frequency of specific mutations can vary dramatically between the molecular subclasses, suggesting that each has its own set of cooperating drivers [39]. However, the specific mutations and epigenetic alterations that define these molecular subgroups are still being determined. Some mutations, such as those in APC and SMAD family member 4 (SMAD4), are common among all the molecular subgroups - suggesting a central role in CRC in general - whereas others are restricted to one subgroup (for example, BRAF in CIMP CRCs) [40].

In CRC, substantial heterogeneity in the specific mutations is evident between tumors, although the mutations seem to cluster in epistatically related groups [41]. The most common mutations in CRC include those in *APC*, catenin-β1 (CTNNB1), KRAS, BRAF, SMAD4, transforming growth factor-β receptor 2 (TGFBR2), TP53, phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit-α (PIK3CA), AT-rich interactive domain 1A (ARID1A), SRY (sex-determining region Y) box 9 (SOX9), family with sequence similarity 123B (FAM123B; also known as AMER1) and ERBB2, which promote tumorigenesis by perturbing the function of key signalling pathways, including the WNT-β-catenin, epidermal growth factor (EGF)-

mitogen-activated protein kinase (MAPK), phosphatidylinositol 3-kinase (PI3K) and TGFβ signalling pathways, or by affecting genes that regulate the central behaviours of cells, such as DNA repair and proliferation [42,43] (Table 1). CRC is frequently initiated by mutations that affect the WNT signalling pathway, and the following neoplastic cells then progress upon deregulation of other signalling pathways, including the RAS-RAF-MAPK, TGFβ, and PI3K-AKT pathways [41,44].

Gene or biomarker	Chromosome	Function	Molecular lesion	Frequency (%)	Predictive?	Prognostic?	Diagnostic?
Tumour suppressors							
APC	5	Regulates the WNT signalling pathway	Inactivating mutations	40–70	No	No	Familial adenomatous polyposis
ARID1A	1	Member of the SWI/SNF family, and regulates chromatin structure and gene transcription	Inactivating mutations	15	No	No	N/A
CTNNB1	3	Regulates the WNT signalling pathway	Activating mutations	1	No	No	No
DCC	18	Netrin receptor; regulates apoptosis, is deleted but not mutated in colorectal cancer, and its role in primary cancer is still unclear	Deletion or LOH	9 (mutation); 70 (LOH)	No	Possible	No
FAM123B	X	Involved in the WNT signalling pathway	Inactivating mutations	10	No	No	No
FBXW7	4	Regulates proteasome-mediated protein degradation	Inactivating mutations	20	No	No	No
PTEN	10	Regulates the PI3K–AKT pathway	Inactivating mutations and loss of protein (assessed by immunohisto-chemistry)	10 (mutation); 30 (loss of expression)	Possible	No	Cowden syndrome‡
RET	10	Regulates the GDNF signalling pathway	Inactivating mutations and aberrant DNA methylation	7 (mutation); 60 (methylation)	No	No	No
SMAD4	18	Regulates the TGFβ and BMP pathways	Inactivating mutations and deletion	25	Possible	Possible	Juvenile polypsis
TGFBR2	3	Regulates the TGFβ pathway	Inactivating mutations	20	No	No	No
TP53	17	Regulates the expression of target genes involved in cell cycle progression, DNA repair and apoptosis	Inactivating mutations	50	Possible	Possible	Li–Fraumeni syndrome
Proto-oncogenes							
BRAF	7	Involved in the MAPK signalling pathway	V600E-activating mutation	8–28	Probable	Probable	Lynch syndrome
ERBB2	17	Involved in the EGF–MAPK signalling pathway	Amplification	35	No	No	No
GNAS	20	Regulates G protein signalling	Mutation	20	No	No	No
IGF2	11	Regulates the IGF signalling pathway	Copy number gain and loss of imprinting	7 (mutation); 10 (methylation)	No	No	No
KRAS	12	Regulates intracellular signalling via the MAPK pathway	Activating mutations in codons 12 or 13 but rarely in codons 61, 117 and 146	40	Yes	Possible	N/A
MYC	8	Regulates proliferation and differentiation	Amplification	2 (mutation); 10 (CNV gain)	No	No	No
NRAS	1	Regulates the MAPK pathway	Mutation in codons 12 or 13	2	Yes	No	No

Gene or biomarker	Chromosome	Function	Molecular lesion	Frequency (%)	Predictive?	Prognostic?	Diagnostic?
Proto-oncogenes (Cont.)							
PIK3CA	3	Regulates the PI3K–AKT pathway	Mutations in the kinase (exon 20) and helical (exon 9) domains	20	Probable	Possible	No
RSPO2 and *RSPO3*	8 and 6, respectively	Ligands for LGR family receptors, and activate the WNT signalling pathway	Gene fusion and translocation	10	No	No	No
SOX9	17	Regulates apoptosis	Copy number gain	9 (mutation); <5 (CNV gain)	No	No	No
TCF7L2	10	Regulates the WNT signalling pathway	Gene fusion and translocation	10	No	No	No
Other molecular alterations							
Chromosome instability	N/A	N/A	Aneuploidy	70	Probable	Probable	No
CpG island methylator phenotype	N/A	N/A	Methylation of >40% of loci from a selected panel of markers	15	Probable	Probable	No
Microsatellite instability	N/A	N/A	Unstable microsatellite repeats in the consensus panel	15	Probable	Yes	Lynch syndrome
Mismatch-repair genes	N/A	Regulate DNA mismatch repair	Loss of protein (as assessed by immunohisto-chemistry), methylation and inactivating mutations	1–15	Possible	Probable	Lynch syndrome
SEPT9	17	N/A	Methylation	>90	No	No	Serum-based assay for cancer detection
VIM, *NDRG4* and *BMP3*	10, 16 and 4, respectively	N/A	Methylation	75	No	No	Stool-based test for early detection
18qLOH	18	N/A	Deletion of the long arm of chromosome 18	50	Probable	Probable	No

Table 1. Common genetic and epigenetic alterations in colorectal cancer.

Epigenetic alterations commonly occur in polyps and CRCs and seem to cooperate with gene mutations to drive the polyp to cancer progression [45,46]. DNA methylation affects CpG-rich regions (CpG islands), which are often located in the 5′ region of genes and can result in transcriptional silencing through effects on transcription factor binding and changes in the chromatin structure [47]. Modifications in DNA methylation related to the development of cancer (in general) include two fundamental changes: hypermethylation of CpG islands in gene promoters, which can silence tumor suppressor genes, and hypomethylation of repetitive genetic elements, which can lead to genomic instability or oncogene activation [48]. Hypermethylation, such as of the septin 9 (SEPT9) gene promoter, is also used for screening purposes.

Importantly, the frequencies of several of these molecular features vary depending on the primary gut location of the tumor being different the ones arising in the right versus those present at the left colon [49,50]. Some studies support a gradual gradient in change in frequency of the molecular alterations, whereas others suggest a more abrupt dichotomy. This has led to the traditional dichotomy of 'proximal' and 'distal' CRC versus adoption of a continuum model. Both models support the notion that the tumor microenvironment (TME) (the gut microbiota and the inflammatory state of adjacent tissue) modulates the way these mutations affect cancer formation and disease progression. Thus, our current understanding of the pathogenesis of CRC is that the disease results from the accumulation of mutations in genes that then drive the formation of the tumor in the context of tumor-promoting factors derived from the adjacent tissue. This paradigm formed the basis for recent recommendation to determine the in situ immune cell infiltrate of the tumor as a prognostic marker combined with its (standard) TNM stage [51]. In close conjunction with these data, recent research has focused on the role of the gut microbiota in colorectal carcinogenesis. Indeed, studies have shown the enriched presence of fusobacteria [52], particularly in cancers with CIMP status [53], which might be inversely related to the CD3+ T cells in CRCs [54]. Together, these data form a basis for further research into the role of the colon microbiota and colon carcinogenesis [55].

1.2 Risk Factors Predicting Colorectal Cancer Recurrence

Several factors seem to influence survival of CRC patients, among which, recurrence is possibly the most important [56,57]. Gan et al reported that 30% to 50% of patients with treated CRC relapse and die due to the disease [58]. Despite the availability of different types of treatment, the chance of recurrence and metastasis of CRC among patients is high. This means that our knowledge about factors influencing the chance of recurrence is either limited or not applicable. Identification of these factors along with better treatment strategies may help in reducing the probability of recurrence and raising the life expectancy of the patients.

At present, the pathological stage is believed to represent the most important prognostic factor for CRC patients. The TNM staging system, defined by the American Joint Committee on Cancer, is the most widely used and is based on parameters such as depth of invasion of the intestinal wall, infiltration of lymph nodes and the presence of distant disease [59]. Although the classical classification of staging provides prognostic information and guides therapeutic decisions, we do not currently have predictive markers of the response to individual therapy in each patient.

For patients with localized CRC- high-risk stage II and stage III - the standard treatment is surgery followed by adjuvant chemotherapy with combination of regimens that include a fluoropyrimidine plus oxaliplatin based on the multicenter studies MOSAIC, NSABP C-07 and XELOXA [60-62]. However, approximately 20-50% of patients will relapse in spite of initial treatment (despite adequate surgery with or without adjuvant treatment).

For these reasons, in patients diagnosed with early stages it is crucial to improve the identification of prognostic factors as well as to identify predictive markers of response to adjuvant chemotherapy and potential therapeutic strategies to personalize adjuvant treatment [59]. Several studies revealed that early detection of disease progression could increase the survival rates and improve overall survival for relapsed patients. Likewise, an important option could be the detection of the relapse in an asymptomatic stage; it could lead to an effective treatment and a better clinical result. However, the currently used monitoring techniques such as imaging techniques (computed tomography (CT)) and the detection of the CEA tumor marker have failed in this regard. The first randomized clinical trial published in 1995 comparing an intensive follow-up regimen versus the standard approach failed to demonstrate the benefit of CT in early detection, although the scanning techniques have been modified at present [63]. Moreover, the detection of the serum level of CEA can be altered by non-malignant conditions including pancreatitis, tobacco or

IBD. Up to 70% of CRC patients have an elevation in their CEA levels at diagnosis, making it a useful marker to monitor the disease after surgery and during treatment. However, controversial results highlighted the low sensitivity and specificity of the CEA, which makes it questionable biomarker for the early detection of relapses [64-66].

1.3 Classification by Molecular Subtypes

CRC represents a group of molecularly heterogeneous disease. During last years, to try to personalize treatment, many efforts have been done to molecularly classify CRC basing on genomic and epigenomic alterations. Our increasing understanding of the molecular pathological epidemiology of CRC has enabled us to refine their classification from a heterogeneous group of diseases with variable clinical outcomes into characteristic molecular subtypes, a development that will allow precision medicine to be implemented and improve the management of patients with CRC. This will permit us a better understanding of potential personalized therapies based on molecular-specific subtypes [67].

1.3.1 Integrated Molecular Characterization (TCGA Classification)

In 2012, the TCGA research network conducted a comprehensive molecular characterization of 224 cases with CRC and analyzed exome sequences, DNA copy number, promoter methylation, and messenger RNA and microRNA expression patterns [68]. A subset of these samples, represented by 97 cases, was examined by whole-genome sequencing. Tumors with mutation rates of >12 mutations per 106 bases (median number of total mutations, 728), which represented 16% of the total number of cases examined, were designated as hypermutated CRCs, whereas tumors with mutation rates of <8.24 mutations per 10^6 bases (median number of total mutations, 58) were termed as non-hypermutated CRCs (84%). Among the hypermutated CRCs, 75% were enriched for MSI, *MLH1* methylation, and CIMP, whereas the remaining 25% presented with somatic mismatch repair (MMR) gene and polymerase ε (*POLE*) mutations, showing mutation rates of >40 mutations per 10^6 bases.

The non-hypermutated CRCs were enriched for somatic mutations in the *APC* (81%), *TP53* (60%), *KRAS*(43%), *PIK3CA* (18%), *FBXW7* (11%), *SMAD4* (10%), *TCF7L2* (9%), *NRAS* (9%), *FAM123B* (7%), *CTNNB1* (*βcatenin*)(5%), *ACVR1B* (4%),and *SOX9* (4%) genes. *FAM123B* (also known as *WTX*) is an X-linked negative regulator of WNT signaling, and most of its mutations involve loss of function. In hypermutated CRCs, *ACVR2A* (63%), *APC* (51%), *TGFBR2* (51%), *BRAF* (46%), *MSH3* (40%), and *MSH6* (40%) genes were frequent mutation targets. Two genes that were frequently mutated in non-hypermutated CRCs were less frequently mutated in hypermutated CRCs: *TP53* (60% vs. 20%) and *APC* (81% vs. 51%).

All non-hypermutated CRCs were characterized as being microsatellite stable (MSS) and were more frequently associated with somatic copy number alterations (SCNAs), indicating that this group is enriched for chromosomal and sub-chromosomal changes.

The WNT signaling pathway was activated in 93% of non-hypermutated CRCs and 97% of hypermutated CRCs, an activation that involved either *APC* inactivation or *CTNNB1* activation, together with changes in many other genes involved in regulating this pathway, including *FBXW7*, *FAM123B*, *SOX9*, and *TCF7L2*. The TGF-β pathway was deregulated in 27% of non-hypermutated CRCs and 87% of hypermutated CRCs. Nearly all CRCs that were examined displayed dysregulation of MYC transcriptional targets because of the activation of MYC by activated WNT signaling and/or inactivation of the TGF-β pathway, indicating an important role for MYC in colorectal carcinogenesis.

This TCGA study opened the opportunity for several novel therapeutic approaches to CRCs, including the use of WNT signaling and CTNNB1 inhibitors, which were shown to be promising in several studies [69,70]. Moreover, several proteins in the MAPK and (PI3K) pathways, including IGF, IGFR, ERBB2, ERBB3, MEK, AKT, and MTOR, were considered to be potential targets for inhibition [68].

Proteomes of CRCs were also analyzed using CRC cases that were characterized by TCGA in 2012. The integrated proteogenomic analyses that were performed demonstrated a functional context for the observed genetic and epigenetic alterations, with relatively few extending to the protein level [71,72]. The genomic and proteomic approach revealed the importance of chromosome 20q amplification, including *HNF4A*, *TOMM34*, and *SRC*. Although *HNF4A* is a transcription factor that plays a key role in normal gastrointestinal development, there are contradictory studies regarding whether *HNF4A* acts as an oncogene or a tumor suppressor gene in CRC [73]. *TOMM34* is frequently overexpressed in CRCs and is involved in the growth of CRC cells [74], whereas SRC encodes a non-receptor tyrosine kinase that is implicated in colorectal carcinogenesis [75].

1.3.2 CRC Gene Expression Profiling (CMS Classification)

During last years, several studies conducted gene expression profiling for categorizing CRCs into subtypes, although the results showed did not lead to a useful single classification [72, 76-78]. Therefore, members of the Colorectal Cancer Subtyping Consortium decided to combine their genomic datasets comprising 4151 samples, including the TCGA source, to generate a consensus molecular subtyping by applying unsupervised clustering techniques [79]. From this, four CMSs groups were established (Figure 3). CMS1 (MSI immune, 14%) is characterized as showing MSI and immune activation; having a CIMP-positive, SCNA-low, *BRAF* mutant phenotype; and it occurs prevalently in older and female patients in the proximal colon. CMS2 (canonical, 37%) is characterized as exhibiting MSS, chromosomal instability (CIN), and WNT/MYC pathway activation; having a CIMP-negative and SCNA-high phenotype; showing the presence of *APC* and *TP53* mutations; occurring in the distal colon to the rectum; and showing superior survival after a relapse. CMS3 (metabolic, 13%) is characterized as showing MSS, having a CIMP-low and SCNA-intermediate phenotype, showing the presence of *KRAS* and *APC* mutations, and exhibiting an epithelial signature and metabolic dysregulation. CMS4 (mesenchymal, 23%) is characterized as exhibiting MSS, being a CIMP-negative and SCNA-high

phenotype, occurring at advanced stages, and showing poorer overall survival and signatures of TGF-β activation, stromal infiltration, epithelial-mesenchymal transition (EMT) activation, matrix remodeling, and angiogenesis. Samples with mixed features (13%) were found to possibly represent either a transition phenotype or intratumoral heterogeneity. Although this CMS classification system was not able to propose a treatment stratification, this subtyping by large datasets facilitated a better understanding of the broad biological groups comprising CRC.

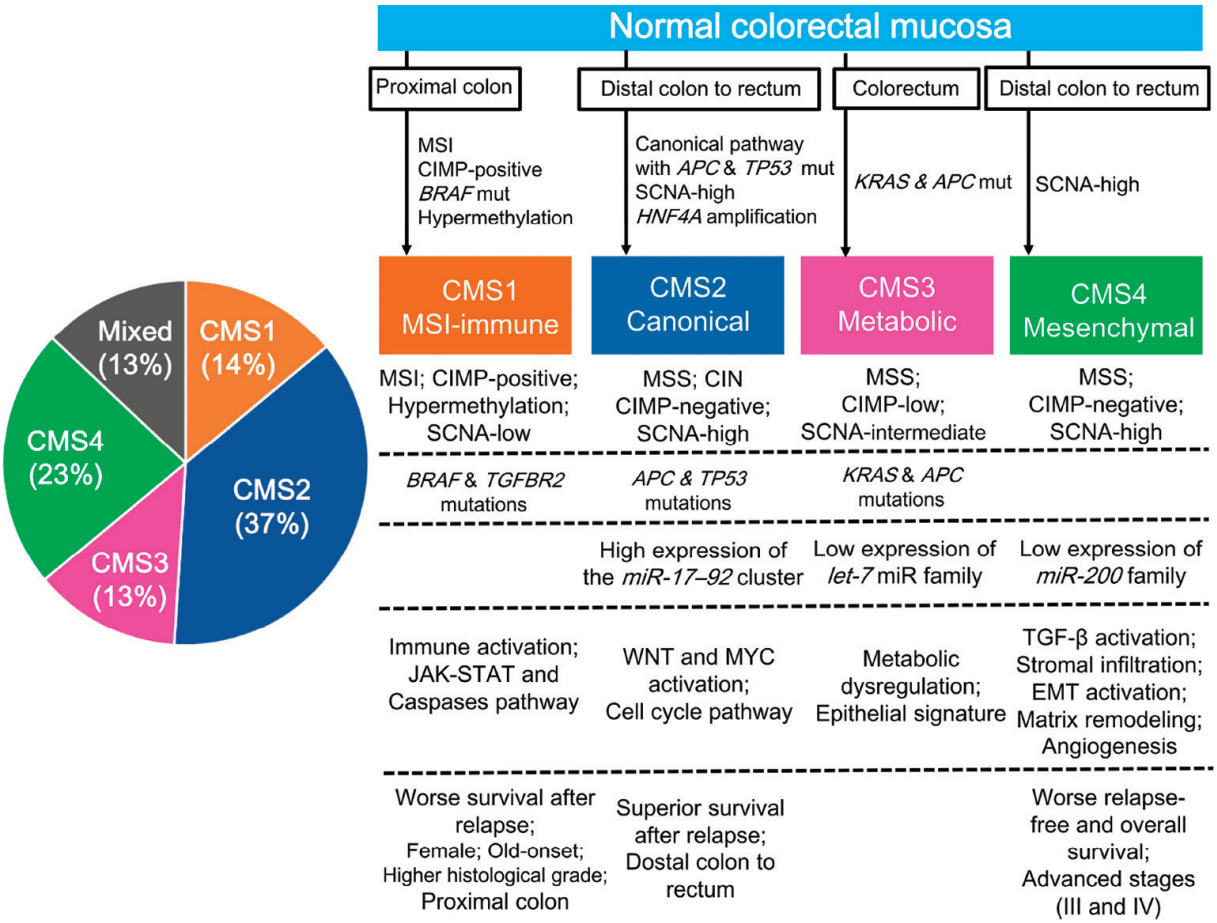


Figure 3. The taxonomy of colorectal cancer according to the Colorectal Cancer Subtyping Consortium, reflecting biological differences in the gene expression-based molecular subtypes [79].

1.3.3 CRC Subtypes Classified by Key Molecular Features

As CRC is a biologically heterogeneous disease, categorization of colon cancers into distinct subtypes using a combination of key molecular features could provide insights regarding the varying clinical outcomes. Using a cohort of patients diagnosed with stage III colon cancer treated in an adjuvant chemotherapy trial, Sinicrope et al. demonstrated that the combination of *KRAS* and $BRAF^{V600E}$ mutations with a DNA MMR status categorized colon cancers into five subtypes with distinct clinicopathological features, including clinical outcomes [80,81]. MMR-proficient tumors with *BRAF* or *KRAS* mutations, comprising 42% of all cases, exhibited higher mortality rates than tumors without this phenotype. *BRAF* wild-type, *KRAS* wild-type, and MMR-proficient tumors, comprising 49%, were the most prevalent subtype in the cohort and were associated with better survival than tumors that lacked this phenotype. Using a population-based cohort of patients with stage I-IV CRCs, Phipps et al. demonstrated that the combination of statuses of MSI and CIMP and mutations of *BRAF* and *KRAS* enabled CRCs to be categorized into five subtypes with distinct clinicopathological features (Figure 4) [82]. Of the five subtypes, type 5 CRCs, comprising 7% of all cases and defined as showing MSI and having a *BRAF* wild-type, *KRAS* wild-type, and CIMP-negative phenotype, showed the lowest mortality rates and were clinicopathologically characterized based on their occurrence in the proximal colon and showed youngest onset. Type 4 CRCs (47%), defined as exhibiting MSS and having a *BRAF* wild-type, *KRAS* wild-type, CIMP-negative phenotype, represented the most prevalent subtype and were clinicopathologically characterized by their canonical pathway with *APC* mutations, and their occurrence in men and in the distal colon to rectum region. Type 4 mostly corresponds to CMS2 (canonical subtype) in the CMS classification [79]. Type 2 CRCs, comprising 4% of cases examined and defined as showing MSS and having a *BRAF* mutant, *KRAS* wild-type, CIMP-positive phenotype, showed the highest mortality rates and were clinicopathologically characterized by their occurrence in females and in the proximal colon and showed late age onset. These two studies suggest that categorization based on key molecular features of CRCs is useful for

understanding the biological features of CRC and for predicting clinical outcomes. Validating the study by Phipps et al. [82], a recent study reported that CIMP positivity could be used to stratify patients with poor prognosis having MSS and *BRAF* mutant CRCs, which correspond to type 2 CRCs in the study by Phipps et al.

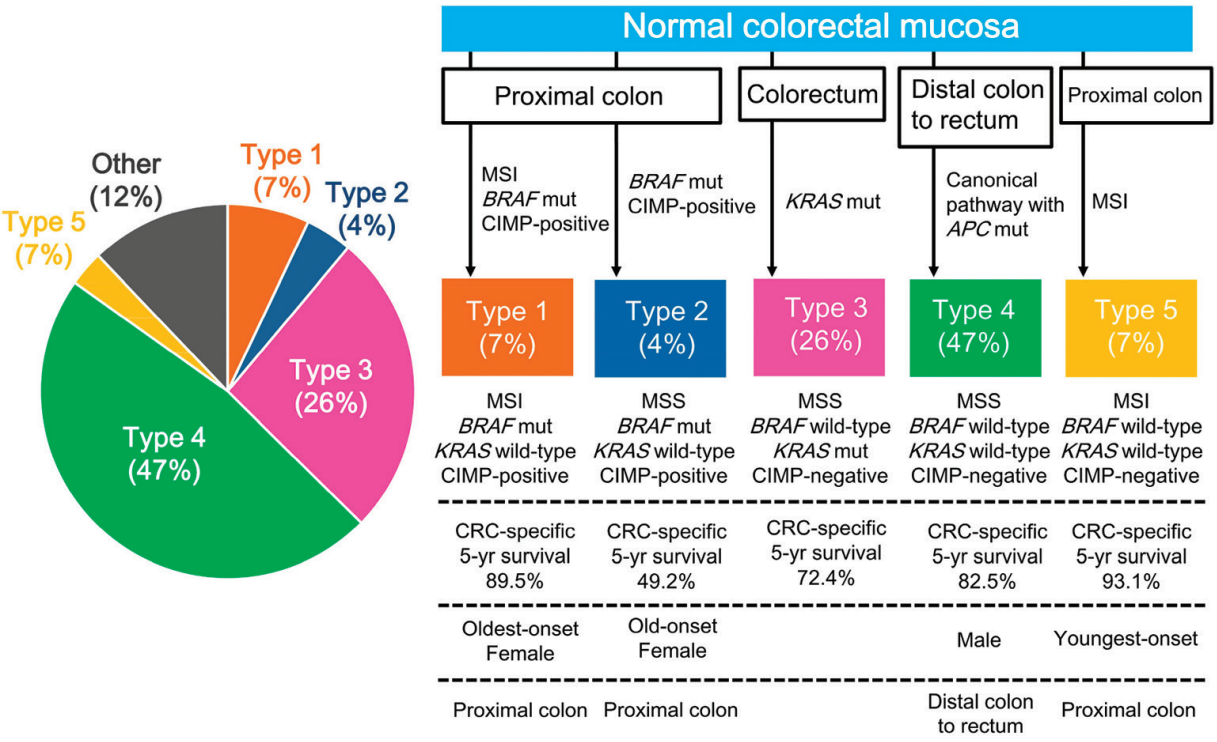


Figure 4. Categorization into five subtypes based on MSI and CIMP status and presence of *BRAF* and *KRAS* mutations [82].

1.4 Molecular Biomarkers

Molecular biomarkers are important for understanding the biological heterogeneity of CRCs and for classifying CRCs into subtypes that can be used to predict prognosis, treatment response, and recurrence risk. These key molecular features or pathways can potentially represent targets for personalized therapies. However, only a few genomic biomarkers, MSI and extended *RAS* and *BRAF* mutational status, and actually *HER2* amplification are routinely used for prognostication and treatment prediction in clinical practice.

1.4.1 CIN

The acquisition of genomic instability is a distinct feature of tumorigenesis, and there are three distinct pathways in colorectal carcinogenesis: CIN, MSI, and CIMP [83]. CIN is a most common feature of CRCs (75-85%) compared with MSI or CIMP [68]. Although substantial progress has been made in identifying the causes of CIN in CRCs, its underlying mechanisms remain unknown [84,85]. Possible mechanisms include alterations in chromosome segregation, telomere dysfunction, and DNA damage response, which affects critical genes such as *TP53* and *APC*. The loss-of-function mutations of *TP53*, which is the main cell cycle checkpoint gene, cause uncontrolled entry in the cell cycle [83]. CRCs with CIN are characterized by the presence of extensive SCNAs throughout the genome and result in aneuploidic tumors and loss of heterozygosity. *APC* mutations, which are associated with defects in chromosomal segregation [86], are also strongly associated with CIN, and thus are likely to lead to CIN and promote cancer progression in CRCs. *APC* forms part of the WNT signaling pathway [80], and its inactivation results in an increase in nuclear CTNNB1 expression and cell proliferation. Thus, the WNT signaling pathway plays a gatekeeper role in CIN CRCs.

1.4.2 MSI

MSI, defined by the National Cancer Institute (NCI) panel markers, BAT26, BAT25, D5S346, D2S123, and D17S250, is a biomarker for defective DNA MMR function in CRCs. According to the classification based on these markers, MSI tumors exhibit instability in two or more markers, whereas MSS tumors show instability in no more than one marker [87]. When CRCs with instability in <30% of markers (MSI-low) were compared with MSS CRCs, MSI-low CRCs did not show any prognostic values compared with MSS CRCs. Therefore, MSI-low CRCs were classified in the same subtype as MSS CRCs [88]. MSI is observed in approximately 15% of sporadic CRCs, most consistently with the frequency of hypermutated CRCs, to which they categorically belong. In addition, frameshift mutations have been detected in the NCI consensus panel of biomarkers; however, multiple other

mutations, including point mutations, also occur in the MMR-defective status. The most prevalent cause of the MMR-defective status in sporadic CRCs is the aberrant biallelic hypermethylation of the DNA MMR gene *MLH1*, which prevents its gene expression. MSI generally results from the inactivation of the MMR genes through aberrant promoter hypermethylation (80% of MSI CRCs; most frequently *MLH1*) or mutations in the MMR genes, comprising *MLH1*, *MSH2*, *MSH6*, and *PMS2* (20% of MSI CRCs). Although CRCs can be analyzed by PCR to detect the presence of MSI, immunohistochemistry can be used to easily evaluate MSI status by demonstrating the absence of a DNA MMR protein. MSI CRCs are mostly enriched for the epigenetic inactivation of the *MLH1* gene, have a CIMP-positive and SCNA-low phenotype, show high frequency of the $BRAF^{V600E}$ mutation and a low frequency of *APC* and *TP53* mutations, and are characterized by their occurrence in females at a late age and in the proximal colon, with poor tumor differentiation and mucinous/signet-ring cell histology. A recent comprehensive molecular characterization revealed that *RNF43* is frequently mutated in CRCs and endometrial cancers [89]. *RNF43* encodes an E3 ubiquitin ligase that negatively regulates WNT signaling. Truncating mutations of *RNF43* are enriched in MSI CRCs and mutually exclusively occur with inactivating *APC* mutations in CRCs. Moreover, an additional study demonstrated the significant co-occurrence of *RNF43* and *BRAF* mutations in the serrated neoplasia pathway [90].

Because nearly all hypermutated CRCs demonstrate a deregulated WNT signaling pathway, this pathway is believed to play a gatekeeper role even in MSI CRCs similar to CIN CRCs. MSI CRCs are clinically characterized as having a favorable prognosis. Furthermore, MSI is a possible marker of sensitivity to therapy with 5-fluorouracil (5-FU). The responsiveness to 5-FU in MSI CRCs seems to depend on the stages of CRCs. Stage II MSI CRCs lack the sensitivity to 5-FU-based adjuvant chemotherapy. In stage III MSI CRCs, the sensitivity to 5-FU-based adjuvant chemotherapy or the standard adjuvant chemotherapy remains controversial, and further studies are required [91,92]. Recent studies suggest that MSI is a marker of good response to 5-FU treatment, particularly in the presence of large deletions in *HSPH1* (*HSP110*) [93].

In May 2017, the US Food and Drug Administration (FDA) granted the accelerated approval to pembrolizumab, a monoclonal anti-PD-1 (PDCD1) antibody, for patients with MSI or MMR-deficient solid tumors. This is the first time that FDA has approved a cancer treatment based on a common biomarker rather than an organ-based approach [94]. MSI causes increased somatic mutations in tumor cells, leading to molecular and biological changes, including high tumor mutational burden, increased expression of neoantigens, and abundant tumor-infiltrating lymphocytes. These changes are associated with an increased sensitivity to checkpoint inhibitor drugs [95,96].

1.4.3 CIMP

Epigenetic instability, which is responsible for CIMP, is another dominant feature of CRCs. The important feature of CIMP-positive tumors is the hypermethylation of promoters of cancer-related genes, which leads to genetic silencing and an absence of protein expression. In CRCs, genetic and epigenetic events are not exclusive, and both cooperate in CRC development, although methylation events are more frequently observed than point mutations. Definitions of CIMP varies substantially among studies with respect to examined foci of methylation and cut-off values for CIMP-positive and CIMP-negative [97]. However, no specific CIMP definitions have been confirmed yet to be superior to the others. A recent study demonstrated that the CIMP status did not show any relationship with CRC prognosis. However, combinations of CIMP with MSI or *BRAF* mutation were associated with CRC survival, although these associations were observed regardless of CIMP status [98]. Because most CIMP-positive CRCs exhibit an MSI phenotype, clinicopathological features of CIMP-positive CRCs overlap with MSI CRCs. Similar to CIMP-positive MSI CRCs, CIMP-positive MSS CRCs are characterized by a high frequency of $BRAF^{V600E}$ mutation; occurrence at a later age, in females, and in the proximal colon, and with poor tumor differentiation. *APC* mutations and activation of the WNT/CTNNB1 signaling pathway are inversely associated with CIMP.

CIMP-positive CRCs arise from a serrated precursor lesion, such as sessile serrated polyp/adenoma [99]. In line with the difficulty of endoscopically detecting a precursor of CIMP-positive CRCs such as sessile serrated polyp/adenoma, CRC diagnosed within five years after colonoscopy is likely to have a CIMP-positive phenotype [100].

1.4.4 POLE mutations

POLE mutations are of increasing interest. These alterations were identified in ultramutated CRCs in the TCGA study [68]. Seven of 30 (23%) hypermutated CRCs lacked MSI, CIMP, and *MLH1* hypermethylation but had somatic mutations in *POLE* and missense or nonsense (but not frameshift) mutations in one or more DNA MMR genes and were designated as ultramutated CRCs [101]. *POLE* encodes one of three polymerases-POLA1, POLD1, and POLE-that are responsible for replicating nuclear DNA and that are involved in the synthesis stage of the DNA repair process; they also play a key role in recombination. Somatic *POLE* mutations apparently cause MSS ultramutated CRCs, unless two DNA MMR alleles of the same gene became mutated by chance. The importance of *POLE* mutations in tumorigenesis has been demonstrated in endometrial cancers and CRCs [102]. Germline mutations of *POLE*, likely to be the cause of predisposition to colorectal and other cancers, and those of *POLD1* map to equivalent sites in the proofreading (exonuclease) domain of *POLE* and *POLD1* and are predicted to cause a correction of mispaired bases inserted during DNA replication. As expected, tumors from carriers of *POLE* and *POLD1* germline mutations are MSS but tend to acquire base substitution mutations [103]. Tumors with *POLE* or *POLD1* mutations are characterized by an extremely high mutation frequency (>1 million per genome) despite MSS. Clinically, only a weak association exists between the presence of mutations in the exonuclease domain of *POLE/PODL1* and increased mortality in MSS CRCs [104]. In contrast to *POLE* and *POLD1*, *POLA1* mutations are rare and functionally impaired because of stringent selection.

Accumulating evidence indicates that MSI and MMR deficiency with high tumor mutational load can predict a response to the anti-checkpoint inhibitors in metastatic CRC (mCRC) [105]. *POLE*-mutant CRC represents an ultramutated but MSS phenotype that is uniquely different from usual CRC with an MSS phenotype. Possibly because of the ultramutated phenotype with a high mutational load and increased expression of neoantigens, patients with treatment refractory mCRC that is characterized by an MSS phenotype and POLE mutations may show clinical responses to pembrolizumab [106].

1.4.5 *RAS*, *BRAF*, and *PIK3CA* Mutations in the MAPK/PIK3 Pathway

The MAPK and PI3K pathways are both involved in cell proliferation, tumor growth and progression. Alterations that affect these pathways contribute to providing proliferative advantages for tumor cells. Mutations of *KRAS*, *BRAF*, and *PIK3CA* are the most common to affect the MAPK/PIK3 pathways in colorectal tumorigenesis. Approximately 40% of CRCs harbor *KRAS* mutations [107]. In contrast, *NRAS* mutations were observed in just 2.5-4.5% of CRCs [108]. *KRAS* and *NRAS* mutations predict resistance to anti-EGFR antibody therapy [109]. In addition, recent studies suggest that *BRAF* and *PIK3CA* mutations also contribute to the resistance to anti-EGFR antibody therapy [110,111]. The prognostic association of *KRAS* mutations in patients with CRC is conflicting. $KRAS^{G12C}$ and $KRAS^{G12V}$ mutations may be independently associated with worse overall survival after diagnosis [112]. One study suggested that the adverse effect of *KRAS* mutations on survival is stronger in distal colon cancers than in proximal colon cancers [113]. In line with this study, *KRAS* or *BRAF* mutations may be associated with shorter overall survival in patients with MSS but not in those with MSI tumors [114]. In contrast, another study demonstrated that the survival of patients with stage II/III CRC might be predicted by CIN and MSI but not by specific driver mutations, including *KRAS*, *NRAS*, *BRAF*, and *PIK3CA* [115].

Approximately 8% of CRCs harbor a point mutation of *BRAF* that is mutually exclusive with *KRAS* mutations. *RAS* mutations are more present in $BRAF^{D594G}$ mutant CRCs than in $BRAF^{V600E}$ mutant CRCs. The $BRAF^{V600E}$ mutation, but not the $BRAF^{D594G}$ mutation, is associated with poor prognosis. More $BRAF^{V600E}$ mutant CRCs were found in the proximal colon compared with $BRAF^{D594G}$ mutant CRCs [116]. As with *RAS* mutations, *BRAF* mutant CRCs are less susceptible to anti-EGFR antibody therapy. In contrast to the pronounced response to $BRAF^{V600E}$ mutant melanoma, vemurafenib, a BRAF inhibitor, does not show a meaningful clinical activity in patients with $BRAF^{V600E}$ mutant CRC. *BRAF* mutant CRCs are resistant to vemurafenib owing to EGFR-mediated re-activation of MAPK signaling [117]. In view of this evidence, the combination of BRAF inhibitor and MAP2K (MEK) inhibitor was applied to *BRAF* mutant CRCs, with modest activity being observed in a subset of patients with metastatic $BRAF^{V600E}$ mutant CRC [118].

PIK3CA mutations are present in 10-20% of CRCs and are associated with other molecular alterations, including the *KRAS* mutant and CIMP-positive phenotype [119]. PIK3CA is an indispensable element of the PIK3 signaling pathway downstream of EGFR. The *PIK3CA* mutation activates the PIK3 signaling pathway, enhancing cell proliferation and eventually leading to carcinogenesis. As with RAS and BRAF mutations, PIK3CA mutations predict resistance to anti-EGFR antibody therapy. CRCs with PIK3CA mutations overexpress PTGS2 (COX2), which plays a critical role in regulating inflammatory responses by generating prostaglandins. Aspirin inhibits PTGS2 expression and downregulates the PIK3 signaling pathway. Regular use of aspirin appears to reduce the risk for PTGS2-overexpressing CRCs but not the risk for PTGS2 weakly expressing or PTGS2-absent CRCs [120]. Furthermore, regular aspirin use after CRC diagnosis is associated with a lower risk for mortality, particularly among individuals with PTGS2-overexpressing CRC. Similar to PTGS2 expression, regular aspirin use was associated with lower mortality in patients with *PIK3CA* mutant CRC but not in those with *PIK3CA* wild-type CRC [121]. A recent population-based cohort study demonstrated that the association of

aspirin use with improved survival differed according to PTGS2 expression but not according to *PIK3CA* mutation status [122]. Another recent study suggests that the regular use of NSAIDs is associated with improved survival in patients with *KRAS* wild-type CRC but not in those with *KRAS* mutant CRC [123]. Genome-wide single-nucleotide polymorphism (SNP) data suggested that the association of aspirin and/or NSAIDs with a lower risk for CRCs differs according to the genetic variation at two SNPs on chromosomes 12 and 15 [124]. The 15-hydroxyprostaglandin dehydrogenase (HPGD) is downregulated in CRCs and functions as a metabolic antagonist of PTGS2. HPGD mRNA expression levels in normal mucosa may serve as a biomarker that predicts a stronger benefit from aspirin chemoprevention [125].

1.4.6 WNT/APC/CTNNB1/TGF-β Pathway

Most sporadic CRCs show abnormal activation of the WNT pathway. Genetic disruption of *APC*, which leads to the activation of the WNT pathway, is a critical early genetic event in colorectal tumorigenesis. In the TCGA study, the WNT pathway was activated in >90% of both non-hypermutated and hypermutated CRCs [68]. Approximately 80% CRCs had *APC* mutations, whereas 5%-10% CRCs exhibited mutations or epigenetic alterations in other WNT signaling components (e.g., *CTNNB1*) that similarly result in the activation of the WNT pathway. APC is not only a critical negative regulator of the WNT pathway but also regulates chromosomal segregation, cellular differentiation, adhesion, migration, and apoptosis.

As a negative regulator of the WNT pathway, *APC* promotes the proteasomal degradation of CTNNB1, which is an important activator of the WNT pathway. If APC is inactivated by mutation, excess cytoplasmic *CTNNB1* accumulates and translocates to the nucleus where *CTNNB1* modulates a transcriptional shift, promoting the activation of *MYC* and many other oncogenes. The disruption of the WNT pathway dysregulates cell proliferation and normal differentiation of colonic epithelia, with adenomas progressing from low grade to high grade owing to the inactivation of other tumor suppressor genes. The transition from adenoma

to invasive carcinoma is usually associated with the inactivation of the *TP53* tumor suppressor gene [83, 126].

The TGF-β pathway plays a critical role in fundamental cellular processes, including cell growth, differentiation, and apoptosis. Chromosomal changes that involve TGF-β strongly contribute to the CIN pathway in colorectal tumorigenesis. The loss of chromosomal 18q is one of the main genomic alterations associated with the inactivation of the TGF-β pathway. Chromosome 18q encodes for two important tumor suppressor genes, *SMAD2* and *SMAD4*, the loss of which inactivates the TGF-β signaling pathway and promotes the evasion of apoptosis and cell proliferation. Nearly all CRCs display MYC activation by the inactivation of the TGF-β pathway and/or activated WNT signaling, indicating an important role for *MYC* in CRC [68]. In the normal colorectal or early CRC tissues, the TGF-β pathway serves as a tumor suppressor by inhibiting cell proliferation and immortalization, and inducing apoptosis; therefore, the inactivation of the TGF-β pathway promotes colorectal tumorigenesis. However, as tumors develop and progress, the tumor-suppressive effects of the TGF-β pathway are often lost. During the late stages of colorectal carcinogenesis, the TGF-β pathway switches to be oncogenic and its activation promotes cancer progression, invasion, and tumor metastasis [127].

1.4.7 TP53 Mutations

TP53 is one of the most important tumor suppressor genes and is the main cell cycle checkpoint regulator [83]. *TP53* inactivation drives tumor progression, allowing excessive cell proliferation. Indeed, the transition from adenoma to invasive carcinoma is usually associated with *TP53* inactivation [83,126]. Loss of 17q, where *TP53* is located, is a frequent event in CRCs because it plays a critical role in the canonical adenoma-adenocarcinoma sequence. *TP53* is more frequently mutated in non-hypermutated CRCs than in hypermutated CRC, similar to the *APC* gene [68]. Of note, not only losses of TP53 activity but also “gain-of-function” *TP53*-mutants mediate tumor metabolic reprogramming, which promotes tumor progression and invasion [128].

1.4.8 Immune Biomarkers

The upcoming evidence of the relevance of immune system in cancer treatment has rapidly revolutionized clinical approach in several solid tumors and also in CRC [129]. It seems that immune checkpoint mechanisms play a critical role in suppressing the anti-tumor T-cell-mediated immune response in the TME. CD274 (PD-L1) is an immune modulator that promotes immunosuppression by binding to PDCD1 (PD-1) of T cells. CD274 engages in the negative regulation of the immune response through the PDCD1 receptor, and evading the host immune surveillance is an important strategy in cancer. Therapeutic antibodies that target PDCD1 and CD274 are effective in numerous malignancies, including CRCs [130]. Tumor CD274 expression is a potential biomarker of a better response to anti-PDCD1/CD274 therapies [131].

Recent studies have suggested the importance of complex associations between tumor molecular characteristics and immune cells in the TME. Emerging evidence has suggested the interactive influences of tumor molecular features with the immune response to the tumor [132]. A recent study demonstrated an association of pro-inflammatory diets such as red and processed meats, with a higher risk for CRC subtypes with absent/low-lymphocytic reaction than CRC subtypes with high-lymphocytic reaction in the TME. The pro-inflammatory diet-associated CRC subtype was enriched in MSS, CIMP-low/negative, and *BRAF* wild-type phenotype. The expression level of CD274 in tumors is inversely associated with the density of FOXP3-positive regulatory T cells, revealing the potential interactions between the immune checkpoint pathway and the host immunity in colorectal carcinogenesis [133]. For a survival association, the association of post-diagnosis aspirin use with better CRC-specific survival seems to be stronger in patients with CD274-low tumors than in those with CD274-high tumors. Moreover, whereas tumors with a high neoantigen load and increased immunogenicity are likely to be a target of immunotherapy, tumors with a high neoantigen load were found to be associated with an increased lymphatic infiltration. In addition, immune cell-infiltrated CRCs were enriched in *HLA* mutations [132].

1.4.9 CDX2

Caudal type homeobox transcription factor 2 (CDX2) expression is highly specific to mature (well-differentiated) intestinal epithelium. Dalerba, et al. hypothesized that lack of CDX2 expression in CRC would indicate a poorly differentiated character (maybe even a highly immature progenitor-cell phenotype) and this would correlate with a worse prognosis [134]. The lack of CDX2 expression defines a small subgroup of CRC, around 7%-13% of cases. In their robust research, using a discovery data-set and an independent validation data-set, the authors confirmed that CDX2-negative tumors were associated with worse prognosis than were CDX2-positive tumors, with lower rates of disease-free survival (DFS) and OS. The negative prognostic impact of CDX2-negativity was also reproducible in the stage II tumors; however, this subgroup was very small in number. In further analysis, they were able to show that administration of adjuvant chemotherapy resulted in improved DFS in CDX2-negative stage II and stage III tumors. CDX2 expression might be an easily accessible prognostic biomarker to define high risk stage II CRC and guide adjuvant treatment in this setting, however prospective validation is needed.

1.4.10 Human epidermal growth factor receptor 2 (*HER2*)

HER2 is an oncogenic driver and is targeted by trastuzumab in the treatment of breast and gastric cancer. Amplification of the *HER2* gene was demonstrated in some CRC xenograft models that are RAS/BRAF wild type and resistant to cetuximab treatment, which renewed the hypothesis that HER2 targeting in this subpopulation might be useful [135]. Around 5% of *KRAS* wild-type CRC patients are found to be *HER2* positive, as defined by the HERACLES Diagnostic Criteria. In an Italian, phase II, prove-of-concept clinical trial (HERACLES), dual-targeted therapy combining trastuzumab and lapatinib was given to treatment refractory mCRC patients who were *KRAS* codon 12/13 wild type and *HER2* positive [136]. In this heavily pretreated subpopulation, 27 eligible patients were identified. At a median follow-up of 94 weeks, 8 patients (30%) had achieved an objective response (among

them 1 with a complete response) and an additional 12 patients (44%) had a stable disease. These results are paving the way for further investigation of this strategy in earlier lines and for trials targeting *HER3* or combined inhibition of *EGFR* and *HER2-4*.

1.5 Novel prognostic tools for recurrence in patients with localized colorectal cancer

Despite current guideline recommendations, clinical decision-making on adjuvant treatment administration remains challenging on the individual patient level, as we are faced with considerable stage-independent outcome variability in daily practice. This underlines the need for validated prognostic biomarkers, telling us which patients are truly at risk of disease relapse, and predictive biomarkers, providing information whether or not this patient will benefit from a certain type of adjuvant treatment or harbors an increased risk of treatment induced toxicity or harm.

Although tumor tissue remains the gold standard for clinical and translational research, there are many barriers in terms of acquisition and usage. To overcome the limitations of tissue biopsies, less invasive techniques are needed to capture the tumor heterogeneity and the molecular changes of tumor cells exposed to therapy [137,138]. The tumor cells release DNA into the bloodstream, and this offers the opportunity to determine the molecular alterations of tumors from the bloodstream, a strategy commonly called “liquid biopsy” [139]. In 1948, Mandel and Métais described for the first time the presence of circulating free nucleic acids (cfDNA) in the bloodstream [140]. In 1977, new studies showed that cfDNA levels were higher in cancer patients than in healthy controls, and we know that the amount of ctDNA is related to tumor burden [141]. Plasma is the best source of ctDNA that exists compared to serum [142] and hosts specific somatic alterations. ctDNA can, in principle, provide the same genetic information as a tumor biopsy.

The bloodstream is a fresh source of DNA so it does not hinder preservation, and obtaining a blood sample through a needle is a minimally invasive procedure and avoids the complications of tumor biopsy.

In addition, blood can be extracted at any time during treatment and allows for dynamic monitoring of the molecular changes that occur in the tumor rather than relying on a static time like a tissue biopsy.

In 2008, Diehl et al. evaluated 162 plasma samples from 18 patients with mCRC demonstrating that ctDNA levels increased with tumor burden, however, whether ctDNA levels are exactly proportional to the systemic tumor burden is unknown. ctDNA was suggested to be a more reliable and sensitive indicator than the current standard marker (CEA) for monitoring tumor burden and the risk of relapse in CRC [143]. In addition, micrometastatic lesions cannot be detected with imaging techniques, which is why these data suggest that ctDNA has the potential to be used as a non-invasive measure of tumor response. For early stages, the levels of detectable ctDNA are probably lower than in metastatic setting. Despite this, ctDNA has been detected in the plasma of multiple types of early-stage cancer. As the sensitivity of DNA analysis has improved, detection of low levels of ctDNA is possible. In a recent study carried out by Bettegowda et al [144] they detected ctDNA in 223 patients with 12 different tumor types in the early stages. This suggests monitoring the occurrence of early relapse after surgery, in order to distinguish between patients who, require adjuvant therapy and patients who are fundamentally cured. In 2015, Garcia-Murillas et al. demonstrated that mutation tracking could identify early breast cancer patients at high risk of relapse. [145]. Recently, Tie et al [146] reported that the detection of ctDNA after surgery in stage II colon cancer provides evidence of minimal residual disease (MRD) and identifies patients with a higher risk of relapse. This could help in guiding the decision of adjuvant treatment in patients with stage II CRC where its role is controversy. The study of the Australian group is limited to stage II patients and lacks a validation cohort as well as validation of the data obtained by NGS, so more studies are necessary to draw solid conclusions.

On the other hand, the monitoring of ctDNA during treatment can help us to evaluate the response to treatment and to detect mechanisms of resistance. It is now clear that the molecular alterations of the tumors are not static but continuously evolving in response to “pressures” (eg, initiation of targeted therapy) [147].

To study clone evolution, a personalized approach would require multiple sequential tissue biopsies. Unfortunately, tissue biopsies only reflect a single point in time of a single site of the tumor implying a risk of losing clinically relevant molecular information. In 2012, Forshew et al. demonstrated that mutations could again be detected non-invasively through re-sequencing plasma samples. They concluded that as chemotherapy regimens stop the growth of other clones, resistance to the EGFR clone, gained dominance in patients with low frequency ovarian cancer. They demonstrated that plasma analysis could identify heterogeneous clones at different sites in the body [148]. On the other hand, TGF-beta and its downstream target IL-6 and IL-11 are cytokines secreted by the primary tumor that facilitate the development of metastases by increasing the survival of the CRC tumor cells [149]. In the clinic, elevated levels of TGF-beta and IL-6 in plasma are associated with a decrease in survival in patients with CRC cancer [150]. In addition, IL-11 has been suggested as a powerful therapeutic target in this field [151]. These results suggest that the detection of circulating cytokines through liquid biopsy could be predictive and facilitate the development of new therapeutic options.

Recently, Cohen JC et al demonstrated that the use of combined assays for genetic alterations and protein biomarkers has the capacity not only to identify the presence of relatively early cancers but also to localize the organ of origin of these cancers [152]. These findings support a multi-omics approach to optimise precision cancer treatment in oncology.

1.5.1 Circulating tumor DNA for detecting Minimal Residual Disease

As previously described, despite curatively intended surgery, a large fraction of patients (30%-50%) experience relapse (4). Identifying the patients at high risk

of relapse, and accordingly those who should receive adjuvant therapy, remains a clinical dilemma. The current approach for estimating relapse risk is based on clinicopathological characteristics, playing TNM staging system the most relevant actor. Unfortunately, this method of evaluation is still far from personalized approach. During last years, several studies have underlined the potential role of minimal residual disease after surgery in determining risk of relapse.

The short half-life of approximately 2 hours for ctDNA indicates that ctDNA is a useful real-time marker of changes in tumor burden (25). However, further work is needed to determine how well the level of ctDNA correlates with tumor burden.

Tie et al. using massively parallel sequencing-based assays evaluated the ability of ctDNA to detect minimal residual disease in plasma samples from a prospective cohort of 230 patients with resected stage II colon cancer. They demonstrated for the first time that detection of ctDNA after stage II colon cancer resection provides direct evidence of residual disease and identifies patients at very high risk of recurrence [146].

Recently, 2 independent groups, Wang [153] and Reinert and colleagues [154] corroborated the poor prognostic implications of ctDNA positivity in patients with stages I to III CRC. Based on findings from these 2-case series, ctDNA assays can be used to classify recurrence risk for patients with nonmetastatic CRC after current standard treatment, and perhaps may be better than the commonly used clinical, laboratory, and pathological risk factors in this setting.

1.5.2 Tumor Microenvironment

Over the last few years, to try answering to the question about which factors could be related to the ability of tumor cells to proliferate and metastasize, the attention has been put on cancer cells molecular characteristics. Recently, due to the improvement in cancer knowledge and to the new evidence of the immune system

as one of the hallmarks of solid tumors, microenvironment has increasingly been recognized as a relevant player in tumorigenesis. For this reason, understanding the different components of TME and the relation between cancer cells and this dynamic structure is an urgent need. The composition of the TME during colorectal tumorigenesis and in advanced cancers is the subject of increasingly intense inquiry and it was demonstrated to be different from normal intestinal stroma. For this reason, evaluating the role of Cancer associated Fibroblasts (CAFs) and cytokines to better understand the paracrine factors that could be related to cancer development and progression is an interesting field to explore. As reported in some previous investigations [155], the majority of genes able to predict cancer recurrence in CRC patients are expressed also in CAFs, independently from cancer cells. Moreover, in a recent series of investigation regarding the immunopofile in localized CRC [156-157], it was showed that immune infiltration was strictly related to patient's prognosis. The presence of the infiltration of particular subsets of T cells in the tumor mass predicts longer DFS intervals after therapy. Additionally, novel CRC classifications based on the type and features of stromal cells will be key to stratifying patients for treatments that target the TME.

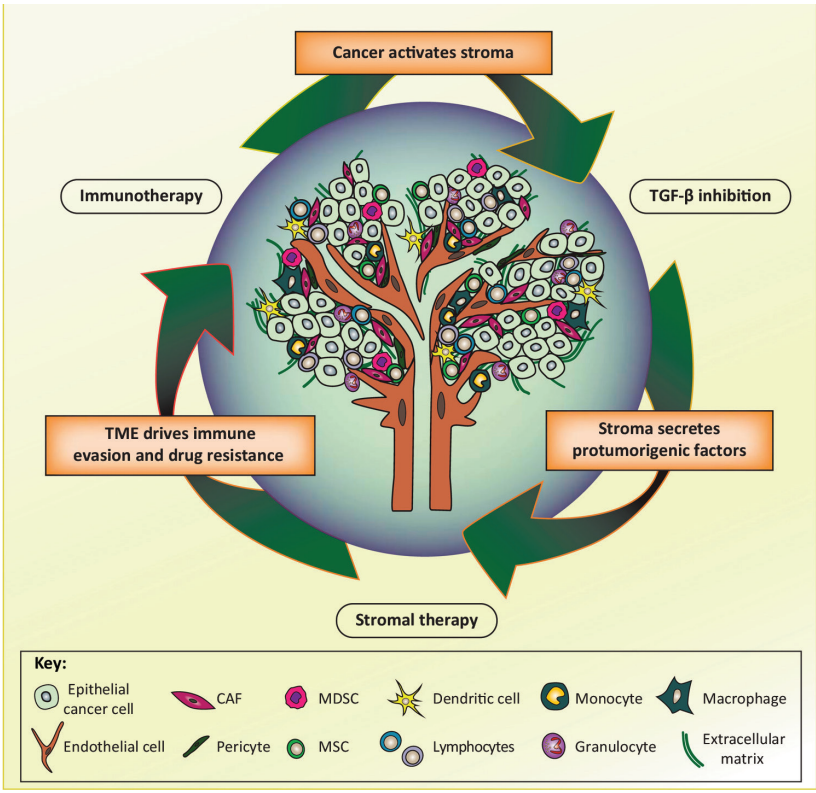


Figure 5. The colorectal cancer ecosystem and the progression-driving role of TGF-b [163].

TGF-β and colorectal cancer

Cytokines are important mediators, connecting the inflammatory TME and cancer cell growth. Cytokines are therefore potential targets in cancer, where their blockade parallel with conventional treatments, allows the targeting of cancer cell extrinsic tumor-promoting pathways. TGF-β activates molecular cascades involved in the control of several biological processes, as cell proliferation, differentiation, migration and apoptosis. TGF-β signalling is initiated by the binding of TGF-β ligands to type II TGF-β receptors (TGFBR2). Three TGF-β isoforms (TGFB1, TGFB2 and TGFB3) are recognized. TGFB1 is the most abundant and ubiquitously expressed isoform. Once bound to TGF-β, TGFBR2 recruits and phosphorylates the type I TGF-β receptor (TGFBR1), which stimulates TGFBR1 protein kinase activity. Activated TGFBR1 then phosphorylates diverse signalling pathways, among them the SMAD complexes. SMAD 4 finally translocates into the nucleus and interact with other transcription factors in a cell-specific manner to regulate the transcription of a multitude of TGF-β-responsive genes [158]. Some of the downstream targets of TGF-β signalling are important cell-cycle checkpoint genes, including CDKN1A (p21), CDKN1B (p27) and CDKN2B (p15) [159]. The role of TGF-β is incredibly interesting due to the capability to be both a suppressor and a promoter of cancer. TGF-β serves as a tumor suppressor in the normal intestinal epithelium by inhibiting cell proliferation and inducing apoptosis. Many CRC escape the tumor-suppressor effects of TGF-β and are resistant to TGF-β-induced growth inhibition. During the late stages of colorectal carcinogenesis, TGF-β acts as a tumor promoter and is usually highly expressed. High levels of TGFB1 in the primary colorectal tumor are described to be associated with advanced stages and a greater likelihood of recurrence and decreased survival [160]. TGF-β has been shown to increase the production of several mitogenic growth factors including TGF-α, FGF and EGF. In addition, TGF-β can activate SMAD-independent pathways, such as Ras/MAPK pathway, JNK pathway and PI3 kinase/Akt pathway. Thus, TGF-β may drive the proliferation of CRC cells in conjunction with these oncogenic pathways. TGF-β is also a potent regulator of cell adhesion, motility and the extracellular matrix

composition, which are involved in tumor invasion and metastasis. In addition, TGF-β signalling promotes angiogenesis and immunosuppression [158]. Therefore, it is likely that cancer cells achieve resistance to the tumor-suppressor effects of TGF-β but remain responsive to the tumor-promoter effects of TGF-β via selective alterations of this signaling pathway.

Whereas genetic and mutational data support a tumor suppressor role for TGF-β signaling in intestinal carcinogenesis, high levels of TGFB1 in the serum of CRC patients associates with poor outcome in the clinical setting [161]. The relevance of TGF-β signaling for disease progression has been widely recognized in tumors where cancer cells retain a functional TGF-β pathway, such as breast or prostate cancer. In these tumor cells, TGF-β induces a variety of prometastatic programmes that range from induction of TME to expression of genes that allow colonization of foreign organs [162]. It is unclear; however, what CRC cells can gain from high TGF-β levels once the pathway is fully inactivated by mutations and how this phenomenon links to an adverse outcome.

In an elegant article recently published [163] the authors speculate that the loss of TGF-β response in the epithelial component of the cancer may facilitate the elevation of TGF-β levels during tumor progression. A functional link between these two events was observed in experimental models. The authors, using a biobank of CRC organoids showed that independence of TGF-β -mediated growth inhibition was associated with advanced stages of the disease [164], whereas loss of SMAD4 in epithelial CRC cells correlated with poor prognosis, as described in other several studies [165]. Nevertheless, abrogation of the cytostatic TGF-β response in CRC cells was found not to be exclusively explained by loss-of-function mutation in pathway components, suggesting that there are additional mechanisms of TGF-β resistance [166]. In addition, a proportion of TGF-β-reactive CRCs may respond to the cytokine by undergoing EMT while bypassing the cytostatic effect. Apparently, this effect is relevant to trigger invasion of a particular subset of early lesions

named sessile serrated adenomas [167], which might therefore benefit from anti-TGF-β therapy. Conversely, it remains unclear whether anti-TGF-β therapy would be safe for patients with cancer cells that have an intact TGF-β pathway even if they represent a subclone and are kept in check by its cytostatic effects. Although the TGF-β pathway has emerged as a powerful architect of the pro-metastatic poor-prognosis TME, its effect on individual stromal cell types is pleiotropic and incompletely charted. Accordingly, the authors proposed that, to treat the cancer ecosystem with more integrated approaches, it is necessary to study and understand its complexity in greater detail. As such, manipulating a master regulator such as TGF-β in sufficiently complex model systems of metastatic CRC should not only provide the needed rationale for clinical translation, but would additionally give us relevant parameters to help map the crosstalk in the cancer ecosystem.

When mutation of TGF-β was analyzed, it was possible to find that mutational inactivation of the TGF-β signaling pathway has a key role during CRC progression. This hypothesis was corroborated by the fact that alterations in TGF-β pathway components are first detected in advanced adenomas and affect 40-50% of all CRCs. In mouse models, mutations in the tumor suppressor APC combined with inactivation of TGF-β signaling components in epithelial intestinal cells trigger the development of invasive adenocarcinomas [168]. Restoration of a functional TGF-β pathway in human CRC cells abrogates proliferation and tumorigenicity [169], implying that TGF-β signaling exerts tumor suppressive effects. Hence, it has been proposed that TGF-β imposes a selective pressure during CRC progression, which tumors avert by genetic inactivation of the TGF-β receptors (TGFBR1 and TGFBR2) or of the SMAD intracellular mediators (SMAD4, SMAD2 and SMAD3) [170].

TGF-β is generally accepted to have tumor-suppressive roles in early-stage cancer [171], an effect that is consist¬ent with a high frequency of mutations in TGF-β signaling components in CRC37. Despite restraining tumorigenesis at early stages, high levels of TGF-β in human CRC are associated with poor prognosis,

particularly in patients with locally advanced disease [172]. Recent data may reconcile this paradox. A study demonstrated that gene expression signatures associated with poor outcome in CRC are driven by TGFβ-induced programmes in stromal cells and that TGF-β promotes metastasis by acting on the TME. Indeed, TGF-β can induce IL-11 expression by CAFs, which promotes STAT3 signaling in CRC and increases multi-organ metastasis. Similarly, TGF-β induces fibroblasts to produce extracellular matrix-remodelling enzymes in co-cultures of CRC cells and fibroblasts [173]. Effective control of mouse CRC progression by the TGFβR1 inhibitor LY2157299 makes targeting TGFβ in patients with late-stage CRC a noteworthy therapeutic possibility. Several TGF-β inhibitors as single agents or in combination are actually being tested in patients diagnosed with diverse solid tumors [174].

Interleukin 6 and colorectal cancer

IL-6 is another cytokine recognized to have several relevant roles in cancer such as progression, proliferation, migration and angiogenesis. IL-6 is produced by diverse cell types and it is a cru-cial mediator of inflammation and immunity. The IL-6 protein is 21-28 kDa in size, depending on the extent of glycosylation [175]. It was first identified as a factor capable of promoting B cell development and regulating the acute-phase immune response [176]. Loss of IL-6 signaling reduces the effectiveness of both innate and adaptive immune responses to invading microorganisms and parasites. IL-6 is produced in the tumor by infiltrating immune cells, by the tumor cells themselves, and by stromal cells. Thus, tumor-associated macrophages (TAMs), granulocytes, and fibroblasts, as well as cancer cells, are all primary sources of IL-6. Nuclear factor-κB (NF-κB), hyperactivated in many solid tumors, is a key transcription factor that drives the expression of IL-6 [177]. Hyperactivation of STAT3 in tumor cells also induces the production of IL-6, thus generating a positive-feedback loop. IL-6 signalling is mediated by two different pathways defined as the classic signalling pathway and the trans-signalling pathway. The classic signaling pathway involves binding of IL-6 to IL-6R on the cell surface and the subsequent interaction of this complex with the membrane-spanning protein

IL-6 receptor subunit-β (gp130; also known as IL-6Rβ) to initiate intracellular signaling. In the trans-signaling pathway, IL-6 binds to a secreted form of the IL-6R (sIL-6R), followed by interaction of the IL-6-sIL-6R complex with gp130. Trans-signaling has a key role in the TME, acting to control the recruitment of leukocytes and the inflammatory activation of tumor-associated stromal cells [178]. The receptor for IL-6 exists in two different form: a single membrane-spanning protein of 80 kDa with a short cytoplasmic domain that lacks signalling capacity and a soluble protein, sIL 6R. During classic signalling, IL 6 binds to the IL 6R that lacks signalling capacity expressed by spe¬cific subsets of leukocytes, megakaryocytes, hepatocytes, and certain barrier epithelial cells. Intracellular signaling is initiated by the formation of the heterohexameric complex, consisting of IL 6, IL 6R, and gp130, followed by the recruitment of cellular signaling proteins, including JAKs and STAT3 [179]. When CRC microenvironment was better typified, it was possible to observe that macrophages are the primary source of IL-6, detected at the leading edges of the tumors, and IL 6 levels have been shown to be closely correlated with advanced-stage disease, as indicated by the number of tumor-positive lymph nodes [163]. Importantly, circulating IL 6 levels have been shown to be prognostic indicators of survival as well as predictors of a response to therapy [176] in several different types of cancer. To date, no clinically relevant genomic alterations in the genes encoding IL 6, IL 6R, or gp130 have been detected in the tumor types analysed by TCGA. Epigenetic alterations have a prominent role in aberrant activation of the IL 6/IL 6R/JAK/STAT3 pathway in cancer, and changes in the expression and/or activation of transcription factors might have a prominent role in the elevated expression of IL-6 in cancer. In particular, JAKs mediate the activation of STAT3 in both non-malignant and malignant cells exposed to IL-6. The anti-inflammatory effects of JAK inhibitors, which have been broadly investigated, are largely attributed to inhibition of STAT3 and/or STAT5 activation. Mutations in the genes encoding JAK enzymes seem to be much less common in solid tumors. On the other hand, the role of STAT3 in solid tumors is widely recognised. Aberrantly elevated STAT3 activity has been estimated to occur in >70% of human cancers. Malignancies in

which hyperactivation of STAT3 has been reported include acute myeloid leukaemia, multiple myeloma, and solid tumors of the bladder, bone, breast, brain, cervix, colon, oesophagus, head-and-neck, kidney, liver, lung, ovary, pancreas, prostate, stomach, and uterus [180-181]. The levels of phosphorylated and/or activated STAT3 have been shown to correlate with a poor clinical prognosis in several of these cancers. Exogenous expression of a constitutively active form of STAT3 confers anchorage-independent growth and tumorigenic capacity on fibroblasts, thus demonstrating the oncogenic activity of the STAT3 protein. This protein has a relevant role in the development and/or progression of multiple cancers, including bladder, colorectal, head-and-neck, lung, pancreatic, prostate, and skin cancers. STAT3 also promotes resistance to conventional chemotherapy and radiation therapy as well as to targeted therapies, such as cetuximab [182]. Indeed, activation of STAT3 via a positive-feedback loop constitutes a primary mechanism of resistance in drug-treated, oncogene-addicted cells. The effects of STAT3 activation on the growth of tumor cells are mediated, in large part, by the STAT3 mediated induction of key target genes that regulate cellular proliferation and metabolism, suppression of apoptosis, and responses to hypoxia. Activated STAT3 also induces the expression of VEGF and selected MMPs, which promote angiogenesis and invasiveness and/or metastasis, respectively. In addition, STAT3 binds to the IL6 promoter, generating a positive-feedback loop, leading to increased IL 6 expression. Both VEGF and IL-6 can also have immunosuppressive effects, which might facilitate immune evasion by tumour cells har-bouring hyperactivated STAT3.

Emerging evidence indicates that STAT3 is also hyperactivated in tumor-infiltrating immune cells and might have profound effects on anti-tumor immunity. Hyperactivation of STAT3 in tumours can occur through a variety of mechanisms. Among others, elevated expression of IL-6 and increased stimulation of IL-6R commonly result in hyperactivation of JAK/STAT3 signaling in tumors. Due to the impact that IL-6 is supposed to have in cancer, many drugs are under investigation. Three main approaches to inhibition of IL-6 mediated signaling at

the ligand and/or receptor level are currently in use: directly targeting IL-6 with antibodies, such as siltuximab [183]; targeting the IL-6R with antibodies, such as tocilizumab[184]; and targeting the IL-6-sIL-6R complex using fusion proteins incorporating sgp130. Direct targeting of IL-6 and IL-6 receptors inhibits both classic and trans-signaling, while targeting the IL-6-sIL-6R complex with sgp130 fusion proteins selectively inhibits trans-signaling.

Interleukin 11 and colorectal cancer

IL-11 belongs to the IL-6 cytokine family. While IL-11 was initially characterized as a cytokine promoting platelet production, recent studies have revealed a critical role of IL-11 in the development of epithelial cancers [174]. In breast cancers, elevated expression of IL-11 is correlated with poor clinical outcome [185]. Despite striking activities within the hematopoietic system, some studies have revealed an unprecedented role for IL-11 signaling as a ‘‘gate keeper’’ for the growth of adenomas and possibly more advanced tumors derived from the gastrointestinal mucosa [186].

IL-11 binds to its specific transmembrane receptor, IL-11 receptor alpha (IL-11Ra). It is believed that the IL-11/IL-11Ra dimeric complex interacts in turn with GP130. The formation of this larger order complex initiates signalling through juxta-positioning of the intracellular Janus (JAK) family tyrosine kinases JAK1, JAK2 and TYK2, which are constitutively associated with a proline-rich intracellular domain of GP130, and enables kinase activation in response to trans-phosporylation [187]. A key STAT3-induced target gene encodes the SOCS3 protein, which terminates GP130 signaling. SOCS3 binds to a membrane-proximal phosphotyrosine residue in GP130 to mediate formation of an E3 ligase scaffold with elongin BC and a cullin protein resulting in ubiquitination of the receptor complex and its proteasomal degradation. Besides activation of STAT3, and to a lesser extent also STAT1, engagement of GP130 also triggers signaling through the RAS-RAF-ERK pathway following GP130 association and subsequent JAK-dependent phosphorylation of

the tyrosine phosphatase SHP2/PTPN11. Finally, GP130 has also the capacity to activate the PI3K-AKT-mTORC1 pathway, although in contrast to engagement of the STAT and SHP2/ERK signaling cascades, the former does not require tyrosine phosphorylation of GP130 [188].

In CRC, CAFs appear to be the principal source of IL-11, where IL11 gene expression was activated in response to exposure to tumour cell-derived TGF-β. The latter is consistent with the identification of two AP-1 motifs in the 50 regions of the IL11 gene that are essential for TGFb1-induced transcriptional activation. Meanwhile, additional cis-regulatory elements in the gene promoter comprise, among others, binding sites for SP-1, STAT3, CTF/NF-1 and possibly NFkB thereby confirming a role for IL-11 during inflammatory processes and both autocrine and paracrine enforcement of STAT3-dependent signalling.

IL-11, rather than IL-6, was shown to drive the growth of gastrointestinal cancers in mouse models. In a genetically engineered model of inflammation-associated gastric cancer, IL-11 acted directly on tumor cells to drive STAT3 activation, cellular proliferation and invasion [189].

Several antibodies that bind to and inhibit IL-11 [186], IL-11Ra or GP130 have been developed and shown efficacy in various mouse models. Likewise, antibodies targeting human IL-11, IL-11Ra or GP130 [190] are being developed for preclinical evaluation, on the back of the clinical success of antibodies targeting IL-6 (Siltuximab) or IL-6Ra (Tocilizumab). Indeed, targeting IL-11 or IL-11Ra may prevent unwanted side-effects arising from targeting GP130.

1.5.3 Multi-Omics Analysis

The acquisition of cancer hallmarks requires molecular alterations at multiple levels including genome, epigenome, transcriptome, proteome, and metabolome.

Earlier detection is key to reducing cancer deaths. Integrated multi-omics approach has demonstrated to detect cancer at early-stage. Recently, Cohen and colleagues described a blood test called CancerSEEK that could detect eight common cancer types through assessment of the levels of circulating proteins and mutations in cfDNA. They analyzed 1005 patients with nonmetastatic, clinically detected cancers of the ovary, liver, stomach, pancreas, esophagus, colorectum, lung, or breast. CancerSEEK tests were positive in a median of 70% of the eight cancer types and the specificity was greater than 99% [152]. Likewise, this same group developed a noninvasive blood test for the detection of pancreatic ductal adenocarcinoma. They combined blood tests for *KRAS* gene mutations with carefully thresholded protein biomarkers to determine whether the combination of these markers was superior to any single marker. The cohort tested included 221 patients with resectable pancreatic ductal adenocarcinomas and 182 control patients without known cancer. The use of *KRAS* in conjunction with four thresholded protein biomarkers increased the sensitivity to 64% and a specificity of 99.5% was found [191]. Concluding that this combinatorial approach may prove useful for the earlier detection of many cancer types.

1.6 Aims:

Hypothesis: **The identification of a comprehensive multi-omic liquid biopsy-based signature is able to detect minimal residual disease, to predict recurrence (prognosis) and to define therapeutic strategies (predictive) in localized colorectal cancer after curative surgery.**

1. Design of a targeted gene panel for ctDNA by bioinformatic screening of gene sets published in CRC.

    1.1 Assessment of the custom NGS genomic panel to detect mutations in patients with localized CRC.

2. Assessment and detection of MRD using circulating tumor DNA by droplet digital PCR in localized CRC patients and its correlation with clinical-pathological parameters.

3. Evaluation of the prognostic and predictive value of a plasma proteomic panel.

    3.1. Determination of cytokine levels involved in the development of metastases (IL-6, IL-11, and TGF-beta) in plasma samples.

    3.2. Evaluation of the association between plasma levels of secreted factors with prognosis.

4. Classification of patients according to molecular subtypes (CRCAssigner and Consensus Molecular Subtypes (CMS) using a customized 96-gene assay for nCounter platform (NanoString Technology)

    4.1. Correlation of the molecular subtypes with ctDNA in patients with localized CRC.

5. Integration of a comprehensive molecular assessment to elaborate a valuable prognostic/predictive multi-omic model.

MATERIAL AND METHODS

2.1 Study design

This prospective study recruited patients from Hospital Clínico Universitario de Valencia (Spain) between October 2015 and October 2017. Inclusion criteria were: age ≥ 18, new diagnosis of histologically confirmed CRC scheduled to undergo surgery with curative intent with no radiological evidence of metastatic disease. Patients with a previous malignancy within the last 5 years were excluded. The study was approved by the human research ethics committee, and all participants provided written informed consent.

We analyzed 150 patients diagnosed with localized CRC over 2 years. Blood samples for ctDNA and CEA analysis were collected at 6 to 8 weeks postoperatively, with serial four-monthly blood samples collected for up to 5 years from a subset of patients (Fig 6). The use of adjuvant chemotherapy was at the discretion of the treating clinician, who was blinded to the ctDNA result.

Per-protocol follow-up included three-monthly clinical review and CEA, with six-monthly CT imaging for 2 years. Thereafter, follow-up was according to the participating institutions' standard of care. Serum CEA was measured by the diagnostic laboratory, with CEA concentrations of ≤3.5 mg/milliliter considered normal. All plasma and tumor samples were analyzed at the Biomedical Research Institute INCLIVA.

Eighty-five of 150 patients had serial plasma available to be analyzed.

For quality control for sequencing, we used 2 HapMap cell lines (NA12878 and NA12877) and two commercial DNAs from Horizon Discovery (HD827, HD734). **HD827** - OncoSpan is the largest and most extensive cell line-derived Reference Standard to date, featuring 386 variants across 152 key cancer genes and **HD734** – Tru-Q DNA reference standards are a highly-characterized, biologically-relevant quality control material used to assess the performance of NGS assays that detect somatic mutations.

Collection of clinical data

This study has used study specific electronic CRFs to collect data. Data collection includes but is not be limited to: demographic characteristics, clinicopathological parameters and treatment received including type of surgery and complications related to surgery.

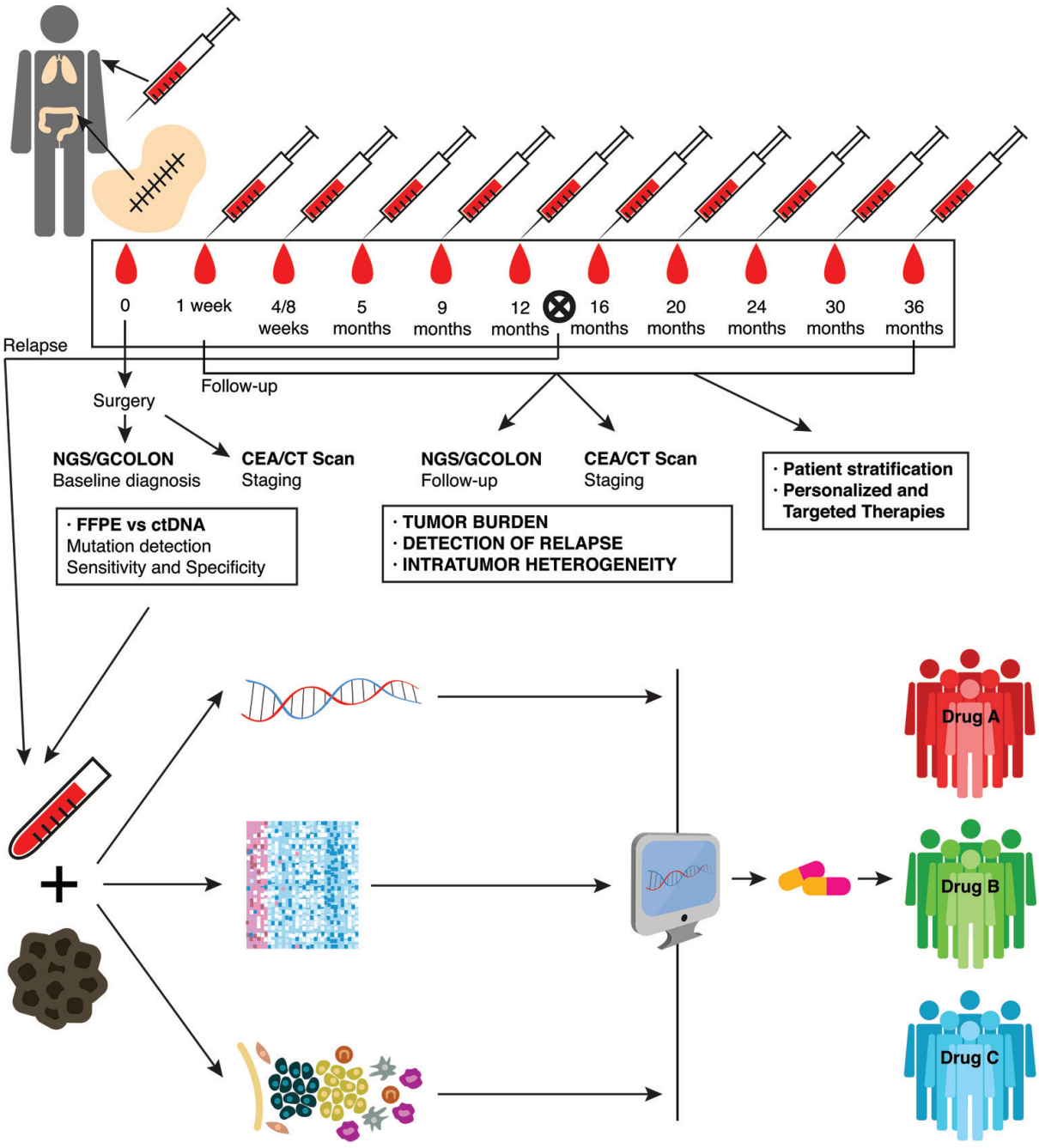


Figure 6. Project strategy-schematic summary

2.2 Patients population

Between October 2013 and October 2015, we enrolled 150 patients and collected 1032 plasma samples from 85 finally eligible patients for the plasma evaluation.

Characteristics		All Patients 150 (%)
Age		70.67 (range: 93-41)
Gender	Man	94 (62.7)
	Female	56 (37.3)
Tumor side	Left	85 (56.7)
	Right	65 (43.3)
Grading	Well	137 (93.2)
	Poor	10 (6.8)
T stage	T1	15 (10)
	T2	17 (11.4)
	T3	91 (61.1)
	T4	26 (17.5)
N stage	N0	94 (63.1)
	N1	39 (26.2)
	N2	16 (10.8)
Stage	I	29 (19.3)
	II	66 (44)
	III	54 (36)
Lymphovascular invasion	Yes	38 (25.9)
	No	109 (74.1)
Perineural invasion	Yes	30 (20.4)
	No	117 (79.6)
CEA level at baseline (ng/mL)	3,5	83 (56.5)
	> 3,5	64 (43.5)
MMR status	Proficient	133 (89.4)
	Deficient	16 (10.6)
Clinicopathologic risk group (Stage II)	Low	38 (58.5)
	High	27 (41.5)
Adjuvant Chemotherapy	Capecitabine monotherapy	34 (22.5)
	Capecitabine + oxaliplatin	21 (13.9)
	No	95 (63.3)
Relapse	Stage I	0
	Stage II	4 (22.2)
	Stage III	14 (77.8)

Table 2. Patient and tumor characteristics

The median age was 70.7 years, and 62.7% of them were male. On the basis of pathological characteristics, 43.3% had the primary tumor in right-sided, 17.5% had T4 disease, 25.9% and 20.4% had vascular and perineural invasion respectively, and 16% had dMMR tumors. Of the 150 evaluated subjects, 55 (36.7%) patients received adjuvant chemotherapy at their clinician's discretion. In all, 44% of patients were stage II and 36% stage III. None of the patients diagnosed at stage I relapsed. In contrast, fourteen (77.8%) and four (22.2%) patients diagnosed with CRC stage III suffered a relapse. The median follow-up was 24.7 (rango 1-45.2) months.

2.3 Pathology

Pathology reports were reviewed for tumor site, LN yield, tumor differentiation, T stage, and LVI. MMR status was assessed by immunohistochemistry for MLH1, MSH2, MSH6, and PMS2 proteins using standard protocols.

We defined a clinicopathologic high-risk group for stage II samples using standard criteria, including pMMR tumors with at least one of the following poor prognostic features: T4, LN yield <12, LVI, and poor tumor differentiation. Paraffin-embedded tumor sections were deparaffinized and the antigens were retrieved in a high pH citrate buffer in a pressure cooker (Dako). The immunohistochemical studies were done using an automated Autostainer Link 48 system (Dako, Glostrup, Denmark) with the following monoclonal antibodies: MLH1 (clone IR079, dilution 1:100. Dako), MSH2 (clone IR085, dilution 1:100. Dako), MSH6 (clone IR086, dilution 1:100. Dako), and PMS2 (clone IR087, dilution 1:100, Dako Only the complete loss of nuclear staining in combination with a positive internal control was accepted as the loss of mismatch repair (MMR) protein expression.). CDX2 immunostaining was carried out with a primary mouse anti-human CDX2 monoclonal antibody (clone DAK-CDX2, prediluted, Dako).

2.4 Custom gene panel design

For the design of the panel content, the primary aim was to allow for the detection of actionable mutations for selection of CRC patients for anti-EGFR

monoclonal antibodies and anti-BRAF therapy and to replace the conventional methods. In addition, the panel content should allow to detect other potential targets for therapies that might become available in the future or mutations with diagnostic or prognostic significance. Our goal is to use this custom panel as a panel for diagnostic molecular profiling of CRC and to monitor ctDNA.

Based on the gene content of commercial cancer panels, the available literature, databases including My Cancer Genome (http://www.mycancergenome.org/) and PCT MD Anderson (https://pct.mdanderson.org/#/) and information retrieved from COSMIC, we developed a custom panel to screen hotspots and complete genes with clinically relevant mutations in CRC. The final content consisted of a target panel of 29 genes including ACVR2A, *AKT1*, AMER1, *APC*, ARID1A, *BRAF, CTNNB1*, *EGFR*, *ERBB3 (HER3)*, ERBB4 (HER4), FAT4, FLNA, *FBXW7*, HRAS, KMT2C, *KRAS*, MEK1, NRAS, *PIK3CA*, POLE, PPP2R1A, *PTEN*, *RNF43*, *SMAD2*, *SMAD4*, *SOX9*, *TCF7L2*, *TGFBR2* and *TP53*.

Canonical transcript ID numbers to evaluate variant effect are provided in **Table 3**. The probes for this targeted custom panel were designed with QIAseq Panel design Tool (Qiagen) and consisted of 1385 amplicons with an average size of 150 bp and a cumulative targeted region of 105 kb.

Gene	*Ts*	*NM-ID*
ACVR2A		NM_001278579
AKT1		NM_005163
AMER1		NM_152424
APC	Y	NM_000038
ARID1A		NM_006015
BRAF		NM_004333
CTNNB1	Y	NM_001098210
EGFR		NM_005228
ERBB3		NM_001982
ERBB4		NM_005235
FAT4		NM_024582
FLNA		NM_001110556
FBXW7	Y	NM_033632
HRAS		NM_176795
KMT2C		NM_170606
KRAS		NM_033360
MEK1		NM_002755
NRAS		NM_002524
PIK3CA		NM_006218
POLE		NM_006231
PPP2R1A		NM_014225
PTEN	Y	NM_000314
RNF43		NM_017763
SMAD2	Y	NM_005901
SMAD4	Y	NM_005359
SOX9		NM_000346
TCF7L2		NM_001198525
TGFBR2		NM_003242
TP53	Y	NM_000546

Table 3. Content of the custom solid targeted tumor panel. Gene names are provided with their classification as a tumor suppressor gene (Ts) or not, and the COSMIC transcript ID used here.

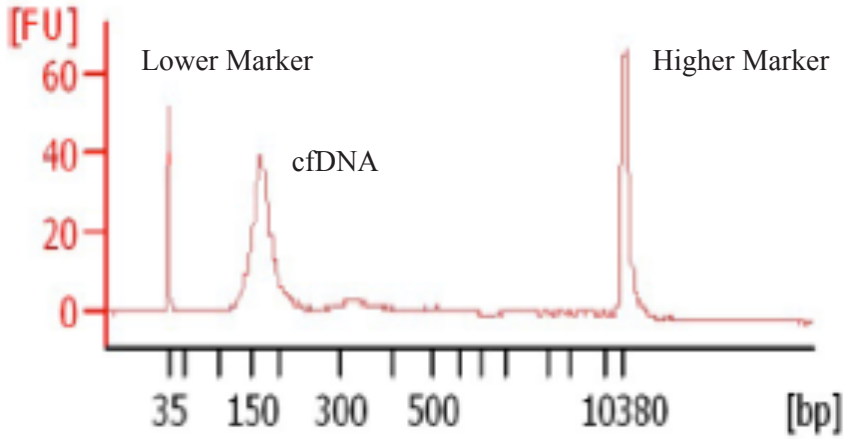
2.5 Processing and DNA/RNA extraction from tumor tissue

One hundred and fifty patients in our cohort had tumor biopsies contemporaneous with the plasma samples. Archival formalin-fixed and paraffin-embedded (FFPE) tissue blocks were retrieved, and thick sections (8 to 10μm) were cut from these blocks, stained with hematoxylin, and macrodissected by using a scalpel to achieve >70% tumor cell purity. A hematoxylin and eosin-stained slide was used to guide manual macrodissection. Tumor DNA was isolated using the Qiagen AllPrep DNA/RNA FFPE Kit (Qiagen) as per the manufacturer's instructions. DNA was eluted in 50 to 100 ml of ATE buffer, depending on the amount of tumor on the slides, and stored at -20°C until quantification.

2.6 Processing of plasma and extraction of circulating DNA

Blood collected in Vacutainer EDTA Blood Collection Tubes (BCT) was processed within 2 hours of sample collection and centrifuged at 1600g for 20 min to separate the plasma from the peripheral blood cells. Plasma was stored at -80°C until DNA extraction. DNA was extracted from 4 ml of aliquots of plasma using the QIAamp Circulating Nucleic Acid Kit (Qiagen) according to the manufacturer's instructions. The DNA was eluted into 70 ml of buffer AVE and stored at -20°C until quantification. For a subset of patients, to test concordance and alternative methods of blood processing, blood was also placed in Streck Cell-Free DNA BCT tubes at the same visit as the EDTA sample, and shipped at room temperature, with plasma separated 48 to 72 hours after venesection.

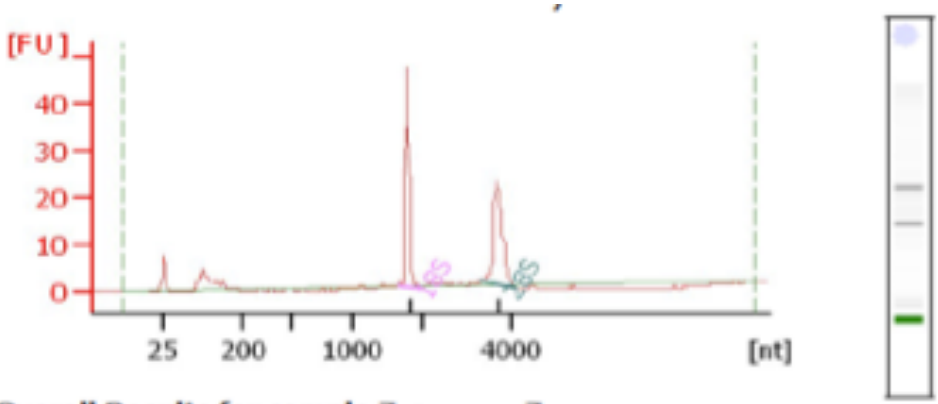
DNA from many plasma samples have been successfully extracted



Plasma patient with appropriate bp, around 166bp mean cell-free DNA (cfDNA)

2.7 Processing and RNA extraction from fresh frozen tissue

Tissue samples from core biopsies taken at surgery were embedded in optimal cutting temperature compound and stored at -80°C until RNA extraction. We cut 10 sections (10 µm thick) from each specimen in a cryostat at -20ºC. RNA extraction from fresh frozen tissue samples were made using the mirVana™ miRNA Isolation Kit (ambion by life technologies), with phenol as per the manufacturer's instructions. The kit employs an organic extraction followed by immobilization of RNA on glass-fiber filters to purify either total RNA, or RNA enriched for small species, from cells or tissue samples.



Overall Results for sample 7 : 7

RNA Area:	103,5
RNA Concentration:	98 ng/µl
rRNA Ratio [28s / 18s]:	1,0
RNA Integrity Number (RIN):	9.6 (B.02.08)
Result Flagging Color:	
Result Flagging Label:	RIN: 9.60

Bioanalyzer: Excellent quality of our samples (**RIN** between 7.5 to 9.6)

2.8 DNA and RNA quantifications from tissue and/or plasma

DNA isolated from tissue or plasma was quantified on a Qubit 3.0 fluorometer (Life Technologies). A Qubit dsDNA BR (broad range, 2 to 1000 ng) Assay Kit and Qubit dsDNA HS (high sensitivity, 0.2 to 100 ng) Assay Kit was used with a Qubit 3.0 fluorometer according to the manufacturer's protocols; a sample volume of 1 μl will be added to 199 μl of a Qubit working solution.

RNA isolated from tissue was quantified on a Qubit 3.0 fluorometer (Life Technologies). A Qubit dsRNA BR (broad range, 20 to 1000 ng) Assay Kit and Qubit dsRNA HS (high sensitivity, 5 to 100 ng) Assay Kit was used with a Qubit 3.0 fluorometer according to the manufacturer's protocols; a sample volume of 1 μl will be added to 199 μl of a Qubit working solution.

2.9 Targeted NGS

Target enrichment was performed on FFPE-extracted DNA using the protocol described in the QIAseq Targeted DNA Panel for Illumina Instruments (May 2017; Qiagen). We applied a 125 ng input DNA, briefly, fragmentation, end-repair and A-addition was performed. Subsequently, barcoded adapters including the UMI and sample-specific indices were ligated to the fragments. DNA was purified and free adapters were depleted by magnetic beads. The targeted enrichment was performed with a customized QIAGEN QIAseq Targeted DNA Panel primer designed to amplify all coding regions. Sixteen samples were processed simultaneously. After library preparation, indexing and bead purification, the libraries were quantified by Qubit analysis to assess successful enrichment and amplification. Libraries were then normalized on beads and pooled for sequencing according to the QIAseq Targeted DNA Panel protocol. The pooled libraries were paired-end (2x151) sequenced on a micro flow cell with V2 chemistry on a NextSeq instrument (Illumina).

2.10 Data analysis and classification of variants

Data analysis was performed using an in-house-developed bioinformatics pipeline (Figure 7). Applying this pipeline, read quality was assessed with FASTQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc). Unique molecular identifiers were identified and included in read headers with UMItools [192]. Adaptors were removed from fastq files with a combination of Cutadapt (https://doi.org/10.14806/ej.17.1.200) and PrinSeq [193]. Reads were then filtered with PrinSeq applying an average read quality score of 30. Remaining reads were aligned against the latest human reference genome (GRCh38) with the Burrows-Wheeler Aligner (BWA-MEM algorithm) [194]. Read alignment transformations and sample enrichment assessment were performed with Samtools [195]. Custom primers were clipped with Bamclipper [196]. Duplicate reads were deduplicated applying UMItools. Variant calling was performed using a combination of three different tools: Mutect2 [197], VarDict [198] and SmCounter [199]. Variants were normalized with BCFtools (part of the Samtools) and merged with VCFTools [200]. Variants were annotated with SnpSift [201] and the Variant Effect Predictor tool from Ensembl (release 92) [202].

To assess panel performance, we added in our pipeline a set of methods to calculate different parameters such as percentage of on-target mapping reads, percentage of duplicated reads, average depth of coverage, percentage of covered bases at different read depths.

Pathological somatic variants were identified based on our cancer mutation database that contains data from different resource such as PCT MD Anderson (https://pct.mdanderson.org/#/), commercial panels (Oncomine and Sequenom), relevant literature and our own expertise.

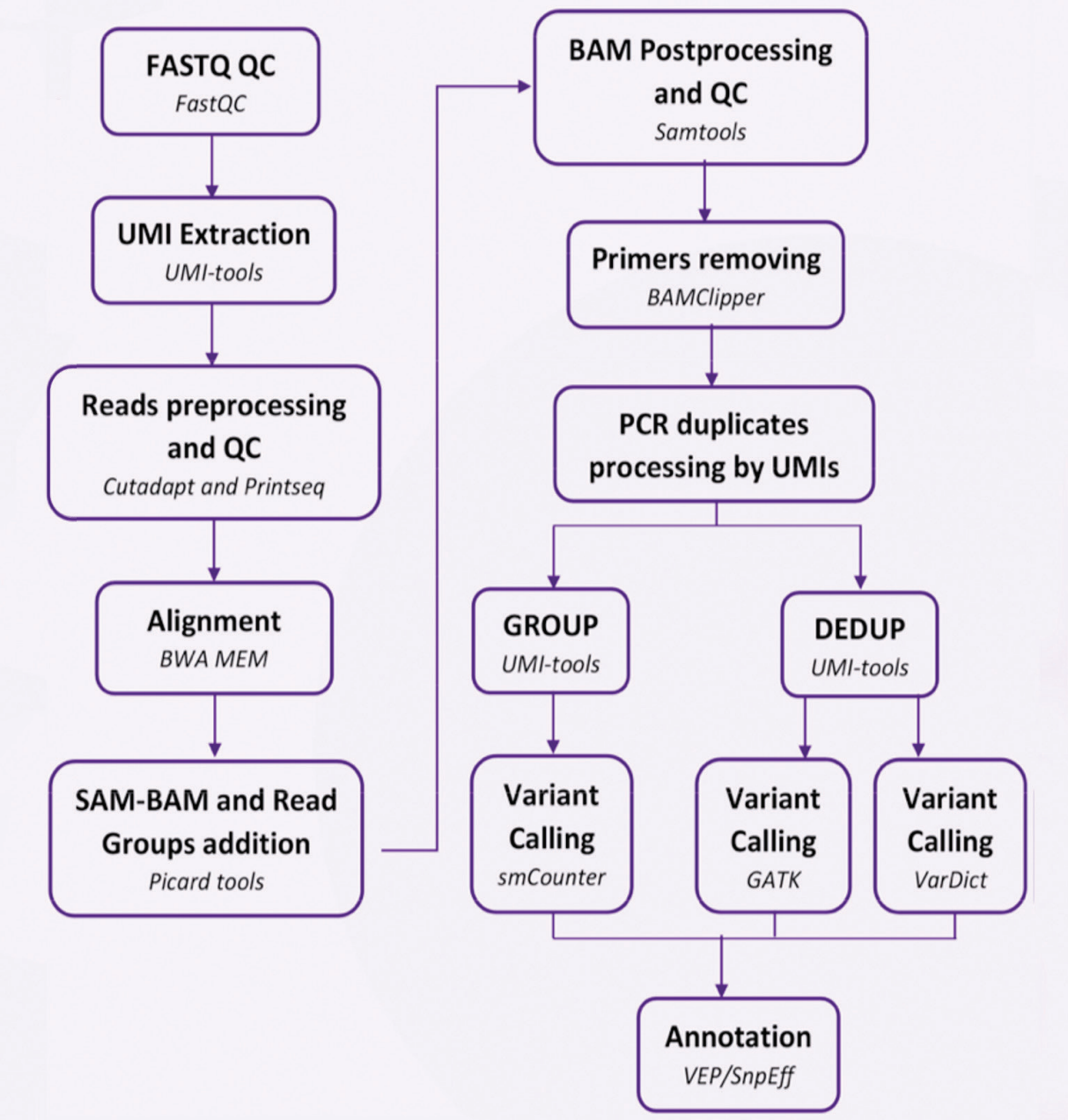


Figure 7. In- a-house bioinformatics pipeline.

2.11 ctDNA analysis

When more than one somatic mutation was identified in a patient's tumor tissue, the two mutations with the highest MAF relative to the MAF in normal control DNA was selected for ctDNA analysis. For each patient, only two mutation identified in the tumor tissue was assessed in the plasma.

Digital PCR analysis

The concentration of mutant DNA (copies of mutant DNA per droplet) was estimated from the Poisson distribution. The number of mutant copies per droplet was calculated as Mmu = −ln (1 − (nmu/n)), where nmu is the number of droplets positive for mutant-FAM probe and n is the total number of droplets. The DNA concentration in the reaction was estimated as follows: MDNAconc = −ln (1 − (nDNAconc/n)), where nDNAconc is the number of droplets positive for mutant- FAM probe and wild type–VIC probe and n is the total number of droplets. The mutation allele fraction = Mmu/MDNAconc. The number of mutant copies per milliliter of plasma was estimated from the mutation allele fraction by taking into account the number of wells run for the sample and the volume equivalent of plasma run, and the mean volume of a droplet (0.89 pl) using the following formula:

Mutant copies per ml = (total number of droplets positive for FAM and/or VIC × 20,000)/total number of droplets read/0.89 × (number of droplets/volume of plasma equivalents)

2.12 ELISA (Enzyme-Linked Immunosorbent Assay): quantitative measurement of IL-6, IL-11 and TGFb in plasma.

The following biomarkers were measured using commercially available ELISA kits according to the respective manufacturer protocols:

TGFβ (ab100647, Abcam®, Cambridge), IL-11 (ab189569, Abcam®, Cambridge) and IL-6 (ab178013, Abcam®, Cambridge). For each assay, samples were measured in duplicate and in-house quality control (QC) samples were included. For each assay, the inter- and intra-assay coefficients of variation were less than 10%. This is consistent with the manufacturer specifications.

Plasma was collected as previously described. ELISA of each cytokines was performed according to the manufacture indication strictly following the instruction

to prepare each solvent and respecting all the times of incubation. Each plate was made up of 96 samples, and 9 controls obtained from randomized man and women who have not a diagnosis of neoplasia were used as control. Each ELISA provided the information about only one cytokine. In this kit specific capture antibodies, each validated for human, are spotted in duplicate on the plate that provide a sturdy base. Briefly, plasma samples were incubated on the well, followed by a biotinylated detection antibody cocktail and chemiluminescence detection method according to the manufactured protocol. Quantification of cytokines levels was determined measuring absorbance at the length of 450 nm λ. The optical density of each sample was compared to a standard curve to determine biomarker concentration. The results were finally analyzed with Excel, and GraphPad v6.0 and validated by R-Studio v1.0.136.

2.13 NanoString nCounter gene expression assays-modified protocol

nCounter® Max Analysis System (NanoString Technologies, Seattle, WA, USA) was used to perform the assay using either standard or modified (Elements chemistry) protocol as per the manufacturer's instructions. For the modified protocol, nCounter ElementsTM TagSets (only capture and reporter tags; NanoString Technologies) and custom-designed target-specific oligonucleotide probe pairs (reporter/Probe A and capture/Probe B probes; from Integrated DNA Technologies, Inc., Leuven, Belgium) were obtained.

The TagSets consist of fluorescently-labelled specific Reporter Tags to resolve and count individual nucleic acid target sequences during data collection (Probe-A), and biotinylated universal Capture Tags to capture the hybridised target nucleic acid sequence to the streptavidin-coated imaging surface (Probe-B). Final hybridising concentration was 20 pM for Probe-A and 100 pM for Probe-B.

For the protocol, 100ng of total mRNA (20 ng/uL) from FFPE tissues was used. Hybridisation reactions were prepared according to manufacturers' instructions

for either 18h at 65°C using Standard CodeSets reagents for standard protocol or for 20h at 67°C using ElementsTM TagSets reagents for the modified protocol. Hybridised samples were pipetted using the nCounter Prep Station and immobilised on to the sample cartridge for data quantification and collection using nCounter Digital Analyzer (NanoString Technologies). The nCounter Prep Station and Digital Analyzer together constitute the nCounter® Max Analysis System. For the PanCancer Progression Panel (NanoString Technologies), the standard protocol from the manufacturer was used. The processing and quality control steps are available in Supplementary Information.

2.14 Statistical analysis

CRCA subtypes were assigned by performing Pearson correlation of gene-wise median-centred expression profiles for each sample with corresponding centroids for the subtypes. The subtype with the highest correlation was then assigned to that sample. Samples were marked as having "undetermined" subtype if the sample's correlation with the subtype centroid had a value (Pearson's r) ≤ 0.15, or if the correlation was high for multiple subtype centroids (Pearson's r difference between first and second highest subtypes ≤ 0.06), in line with the published CMS classifier [78]. CMS subtypes were determined from microarray or RNAseq data using the CMSclassifier R package (v1.0.0) and the classify CMS function, using the single sample prediction (SSP) classifier.

Cytokines evaluation was performed by ELISA assays. Firstly, results were analyzed with a Shaprio-Wilkinson test followed a Kruskal Wallis or T test according to Mann-Whithey correction. P value was indicated as follow: * <0,05, **< 0,01 and ***< 0,001 for the univariate analysis. Results were considered statistically significant when the obtained *p* value was ≤0,05.

A univariate analysis will be performed with a logrank test. Those variables which will achieve statistically significant results will be included into a multivariate

Cox regression analysis. Those variables which value are continue, will be recoded and scored according to the results of their AUC obtained from a ROC curve. To be significant a p value of ≤0,05 was needed. The results were finally analyzed with Excel, and GraphPad v6.0 and validated by R-Studio v1.0.136.

RESULTS AND DISCUSSION

3.1 Design of a targeted gene panel for ctDNA by bioinformatic screening of gene sets published in CRC. (Aim 1)

We wanted to implement an NGS strategy to globally analyze patients with diagnostic quality to enhance our understanding and control we had in pre-NGS strategies. To enable precision oncology in patients with solid tumors, we developed a targeted gene panel, an amplicon-based NGS assay for targeted deep sequencing of selected exons of 29 key cancer genes in FFPE tumors. Barcoded libraries from tumor samples were constructed, sequenced, and subjected to a custom analysis pipeline to identify somatic mutations. To assess the analytical sensitivity of the targeted NGS as a tool for genetic testing, we tested 2 HapMap cell lines (NA12878 and NA12877) and two commercial DNAs from Horizon Discovery (HD827, HD734) samples with previously known point mutations and insertions/deletions in cancer genes. All known variants were accurately detected, and there was high reproducibility. The minimum variant allele frequency (VAF) was set up to 5%. Our validation in control samples demonstrated that our approach identified oncogenic DNA alterations in clinical solid tumor samples with high accuracy and sensitivity.

The following criteria were chosen, based on the absence of false positive or false negative variants in the control dataset, to process each sample:

- Bases with Phred score less than 30 (one incorrect base call in 1000) were rejected.
- Minimal depth of coverage per base >= 100x. Variants in hotspots with a coverage <100 were regarded as non-reliable
- VAF of 5% as the minimal thresholds.

3.1.1 Assessment of the custom NGS genomic panel to detect mutations in patients with localized CRC.

Next, to characterize somatic mutations in CRC, we analyzed 150 FFPE samples from patients diagnosed with localized CRC. Variant calling directly depends on depth of coverage. The higher the coverage, the higher the variant confidence is. The average depth of coverage per sample was 200x. Bases with a depth of coverage less

than 100x were discarded for further analysis. Two samples (1.3% of total samples) had insufficient coverage and thus discarded. A minimum VAF of 5% was applied for variant filtering.

Based on that, we found 18 samples (12%) with a VAF >5% but without known pathogenicity. To summarize, we found a known mutation in 132 patients (88%).

Gene mutations

Overall, we identified 12 somatic recurrently mutated genes (Figure 8). The frequency of mutation in those genes was *TP53* (50%), *APC* (47%), *KRAS* (47%), *PIK3CA* (23.5%), *BRAF* (8.33%), *FBXW7* (6.81%), *NRAS* (6.1%), *CTNNB1* (3.8%), *SMAD4* (3.03%), *PTEN* (3.03%), *ERBB3* (2.3%) and *EGFR* (1.5%). As expected, the mutated *KRAS* and *NRAS* genes usually had oncogenic codon 12 and 13 or codon 61 mutations, whereas the remaining genes had inactivating mutations.

In our cohort, tumors with MSI (10.7%) showed 25% and 18.8% mutation in *BRAF* and *KRAS* genes. 14 out of 16 (87.5%) patients with right-sided primary tumor site. According to primary site, *BRAF (80%)* and *PIK3CA (65.5%)* were frequently found in right-sided CRC. It should be note that TP53 mutations (69.4%) were found in left-sided CRC.

The median and mean genetic variants per patient were 317.0 and 336.4, respectively (Figure 9). We calculated the percentage of variants per gene and it was as follows: *TP53* (25.3%), *APC* (24.2%), *KRAS* (22.4%), *PIK3CA* (11.7%), *BRAF* (3.9%), *FBXW7* (3.2%), *NRAS* (2.8%), CTNNB1 (1.8%), *PTEN* (1.4%), *SMAD4* (1.4%), *ERBB3* (1.1%) and *EGFR* (0.7%).

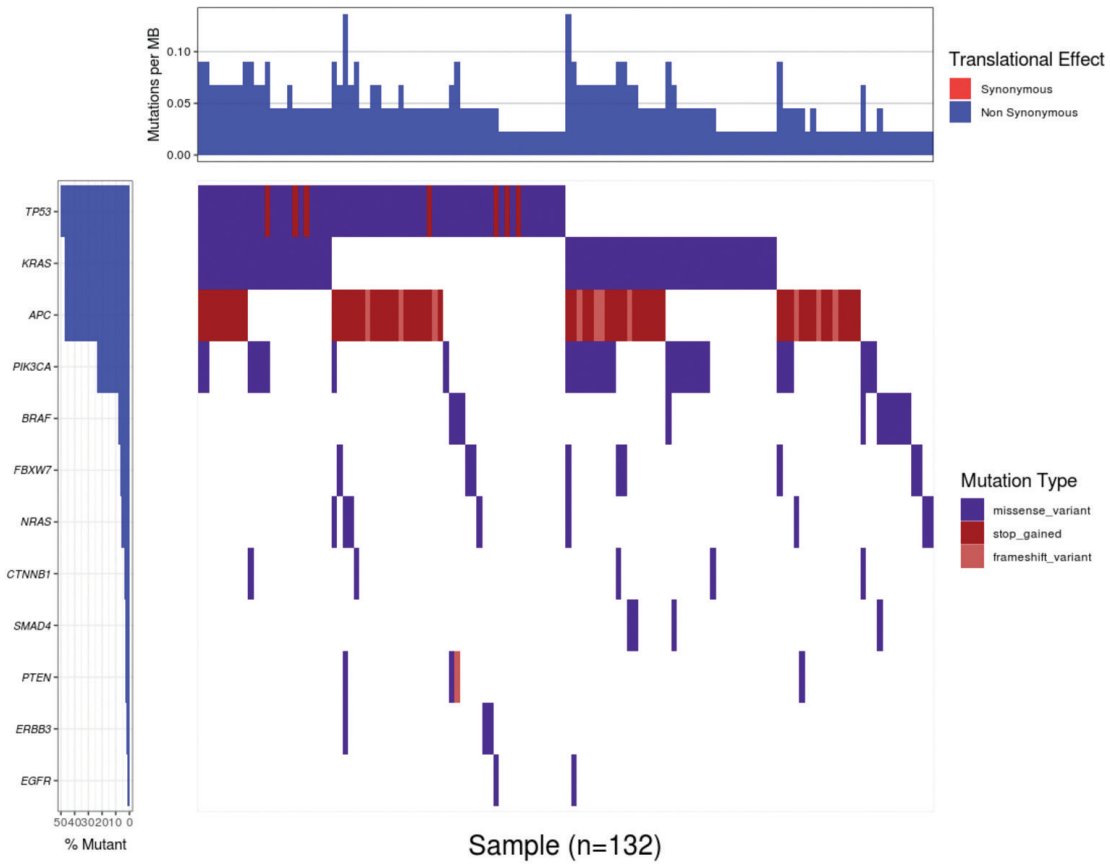


Figure 8. Waterfall: tumor samples grouped by gene mutation rate. Mutation frequencies in each of the tumor samples from 150 patients is illustrated. The percentage of mutations in each gene is shown on the left.

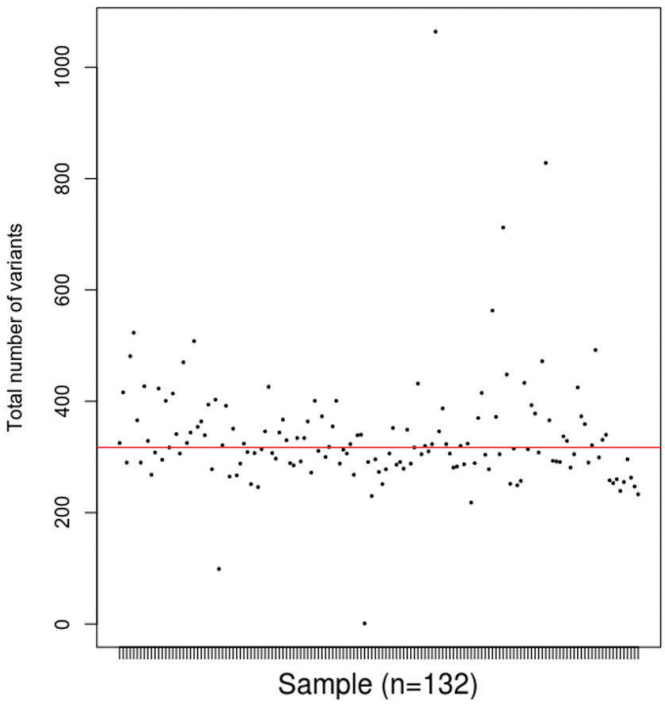
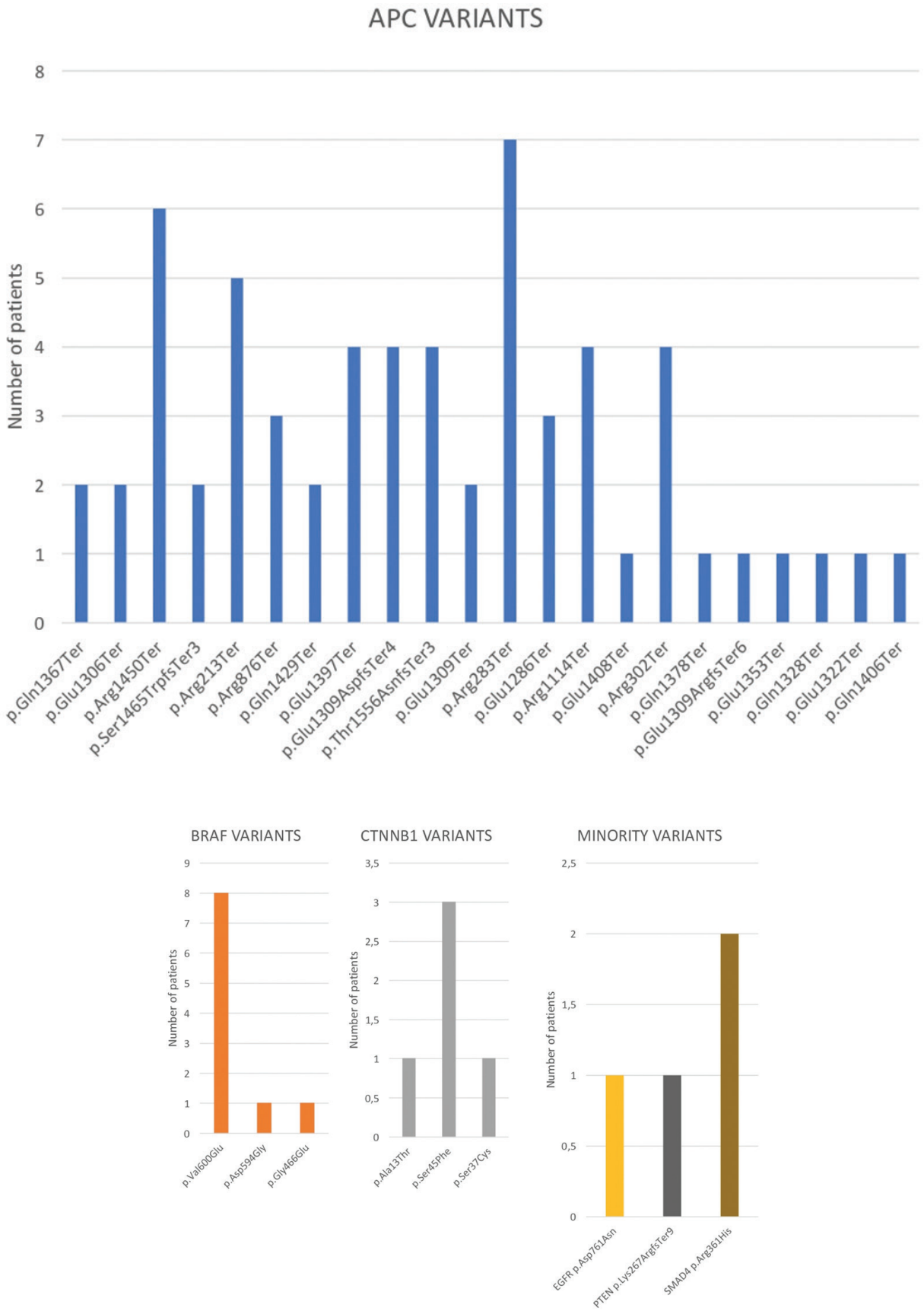


Figure 9. Percentage of variants per CRC patient.

Aditionally, We provide (Figure 10) the types of variants per gene.



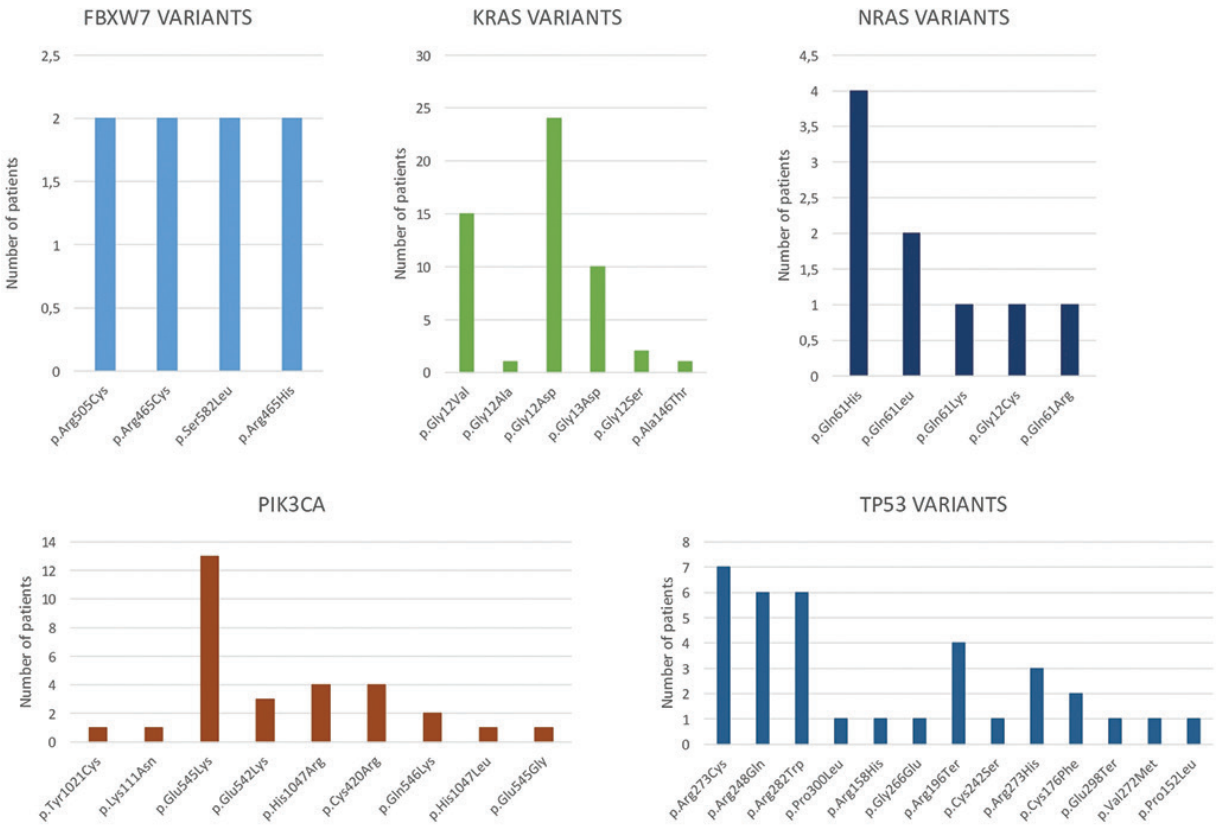


Figure 10. Variants from NGS analysis.

Discussion

Targeted therapies for solid tumors have shown great promise and based on ongoing clinical studies, the potential to enlarge the current set of approved anti-cancer drugs is huge. As most of these drugs target specific signaling pathways it is of outmost importance to detect mutations in the critical genes involved allowing for precision medicine. As the number of actionable genes in solid tumors is steadily growing there is an increasing need to perform multi-gene mutation testing in molecular diagnostics. Several NGS gene panels are commercially available but these panels often contain genes or hotspots that are not of particular interest for molecular diagnostics due to their uncertain clinical significance, or to the lack of tumor types studied. We developed a custom panel to screen hotspots in 29 genes for clinically relevant mutations in CRC. Our selection was based on information retrieved from COSMIC, My Cancer Genome, PCT MD Anderson, and

the available literature. We based our selection on 1) the present or likely future clinical significance in terms of therapeutic, prognostic or diagnostic value, both from clinical research and pre-clinical data and 2) the frequency of known hotspot mutations. We found that the resulting total size of 105 kb is feasible, offering a sufficiently extensive and clinically relevant mutational profiling in a cost-efficient way. We also validated our NGS data through ddPCR (Figure 9) and demonstrated a robust correlation (R^2 =0.841) between both methods. Quality control testing of the Horizon reference samples (accuracy) identified all mutations at about the reported frequencies down to 2%. Previous series confirmed that variants could be detected even below the 3% allele frequency but to be on the safe side, we have set our threshold at 5% with a minimal mean coverage set at 100. Using these thresholds, all expected variants could be correctly demonstrated. This arbitrarily set VAF is used by several diagnostic laboratories. Actually, it also remains to be investigated what the biological or clinical significance is of (subclonal) mutations with a true low VAF, not caused by a low percentage of neoplastic cells in the starting material, which should always be taken into account.

Classification of variants remains a challenging task especially for those variants for which no or restricted information is available in databases, the literature, or on dedicated websites. We found 18 samples (12%) with a VAF >5% but without known pathogenicity. We classified mutations as ‘pathogenic’ when there is sufficient evidence for a role in cancer based on our cancer mutation database that contains data from different resource such as PCT MD Anderson, commercial panels, relevant literature and our own expertise. However, classifying variants in terms of both pathogenicity and clinical relevance often remains a rather subjective task and thus should best be performed in a clinical molecular tumor board where expertise is available from pathologists, oncologists, clinical biologists, bioinformaticians, and NGS scientists [203]. It would be useful, paired analysis of tumors and patient-matched blood samples that would allow unambiguous detection of somatic mutations to guide treatment decisions. Regardless, the community of molecular

diagnostic oncology would benefit significantly from a standard scheme of classes as proposed by others [204], from guidelines how to classify variants, and from a curated classification database. Initial efforts towards these goals have recently been reported including a variant annotation scheme [205] and a database of potentially actionable cancer mutations.

Screening of our set of 150 CRC samples revealed mutations in 88% of the specimens, predominantly in *TP53 (50%)*, *APC (47%)*, *KRAS (47%)*, *PIK3CA (23.5%)*, *BRAF (8.33%)*, *FBXW7 (6.81%)*, and *NRAS (6.1%)*. Overall, our data are in good agreement with those reported by others [206].

3.2 Assessment and detection of MRD using circulating tumor DNA by droplet digital PCR in localized CRC patients and its correlation with clinical-pathological parameters. (Aim 2)

Personalized tumor-specific assays ddPCR for mutation tracking

We investigated the potential utility of ctDNA analysis in localized CRC in a prospectively accrued cohort of 150 patients. We subjected primary tumor DNA, extracted from a tumor biopsy at diagnosis before treatment, to NGS, identifying one or more somatic mutation(s) in 88% [132 of 150] of cancers, with two or more mutations found in 86 (65.2%) cases.

To track these mutations in plasma DNA and identify the presence of ctDNA, we designed personalized ddPCR assays for each somatic mutation identified. ddPCR can accurately quantify mutant DNA at single-molecule sensitivity, even in the presence of vast amounts of wild-type DNA [207]. NGS and ddPCR analysis had a high level of agreement in baseline tumor DNA in the assessment of the mutant allele fractions (Figure 10), demonstrating the robust ability to develop ddPCR assays for diverse mutations. In patients with two mutations identified in the primary tumor, we tracked both mutations in plasma. Thirteen of 65 patients with plasma available post-surgery (20%) harbored the tumor-specific mutation in their plasma (ctDNA-positive) 6 to 8 weeks after surgery, hereafter termed “postoperatively.”

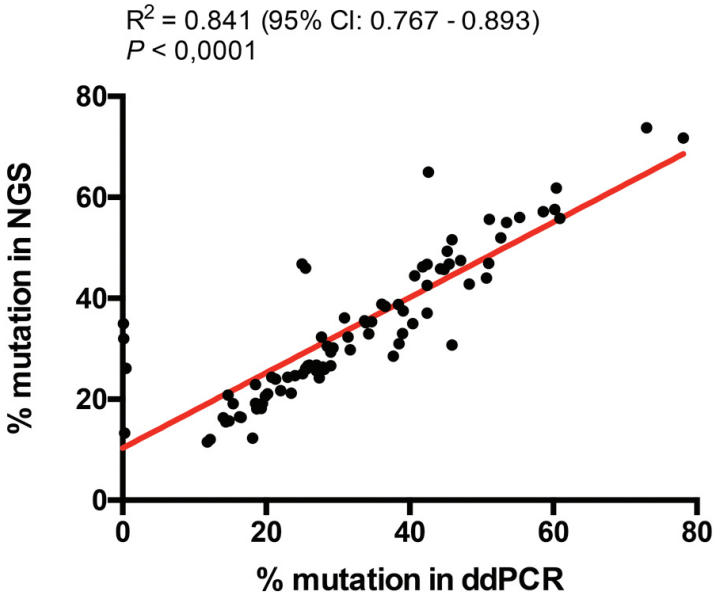


Figure 10. Correlation between NGG and ddPCR.

Of the 85 patients with available plasma samples, 33 (38.8%) patients received adjuvant chemotherapy at their clinician's discretion. These patients were younger and had high-risk features. There was no significant association between postoperative ctDNA status and conventional high-risk clinicopathological factors (Table 4). As of 8 May 2019, median follow-up was 24.7 months (range, 1 to 45.2 months). During this period, 18 (21.2%) patients experienced radiologic recurrences, including 6 of 52 (11.5%) patients not treated with chemotherapy and 12 of 33 (36.4%) patients treated with adjuvant chemotherapy.

Tracking mutations in ctDNA to identify MRD and anticipate relapse

We next assessed whether ddPCR could be used for the detection of ctDNA to predict early relapse. The personalized ddPCR assays were used to track mutations in serial plasma samples taken at baseline, postoperatively with the sample taken 6 to 8 weeks after surgery, and then every 4 months during follow-up. We also assessed the potential to predict relapse from the different time points of ctDNA analysis, starting with the baseline plasma sample taken at diagnosis before any treatment. Consistent with previous observations [145], ctDNA was detected in 61.2% (52 of 85) of baseline plasma samples. ctDNA detection at baseline, before any treatment, was not predictive of DFS. Concentration of ctDNA in pretreatment plasma was significantly lower with stage I compared with those with stage II-III tumors (p=0.036, Figure 11). On the other hand, CEA pre- and postoperative was not predictive of DFS (Figure 12).

We next assessed the potential of a single postsurgical sample taken 6 to 8 weeks after surgery. ctDNA was detected in the single postoperative blood test in 19.7% (13 of 66). Detection of ctDNA in a single postsurgical sample was a significant predictor of early relapse in a multi-variable model (Table 4). Nineteen patients did not have a sample taken at 6 to 8 weeks after surgery and were excluded from analysis of this time point.

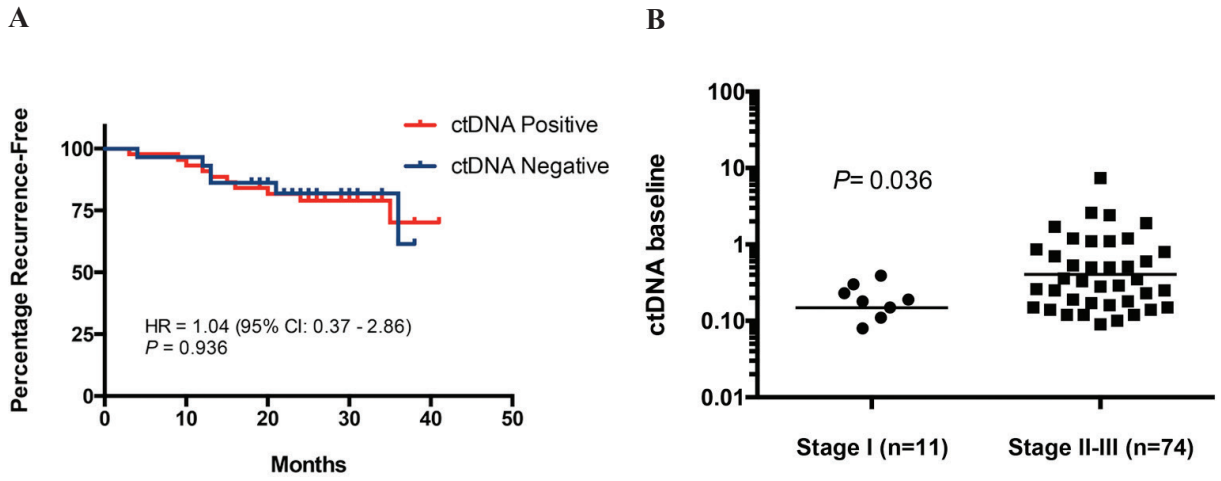


Figure 11. Early relapse is not predicted by analysis of baseline ctDNA. Disease-free survival according to the detection of ctDNA in the baseline plasma sample (A). P value determined by log-rank test. Pretreatment ctDNA concentration according to stage (B).

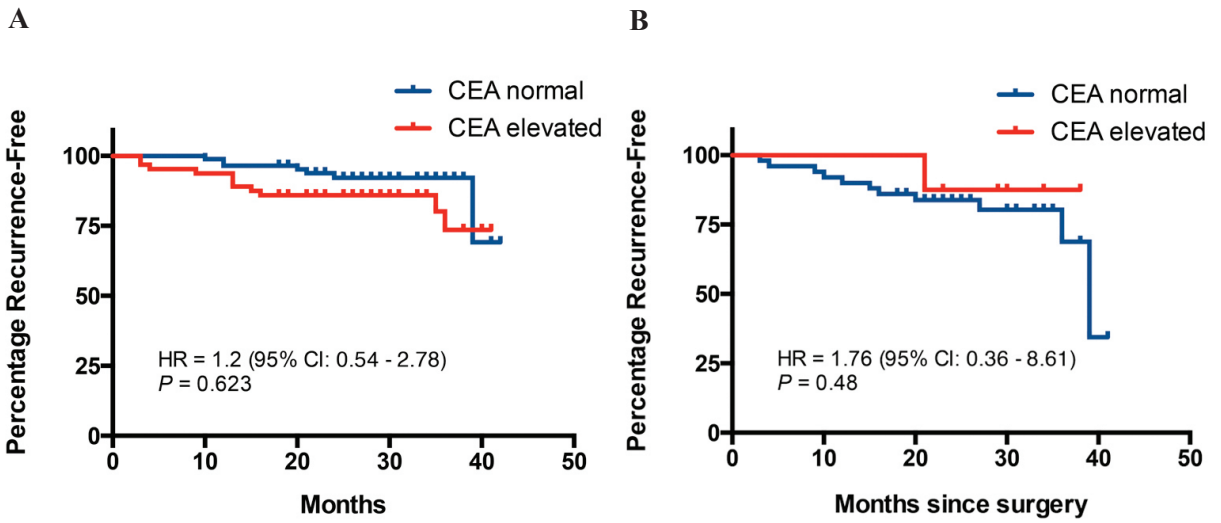


Figure 12. Early recurrence is not predicted by analysis of baseline and postoperative CEA. Disease-free survival according to the level of CEA at baseline (A) and postoperative (B). P value determined by log-rank test.

We examined whether the detection of ctDNA in serial samples, which we termed "mutation tracking," could improve relapse prediction compared with a single postsurgical sample (Figure 13). Detection of ctDNA in serial samples was predictive of early relapse [disease-free survival: median of 38 months (ctDNA detected) versus median not reached (ctDNA not detected); HR, 11.3 (95% CI, 3.72 to 34.52)]. Detection of ctDNA by mutation tracking was a significant predictor of early relapse in a multivariable model [p=0.004].

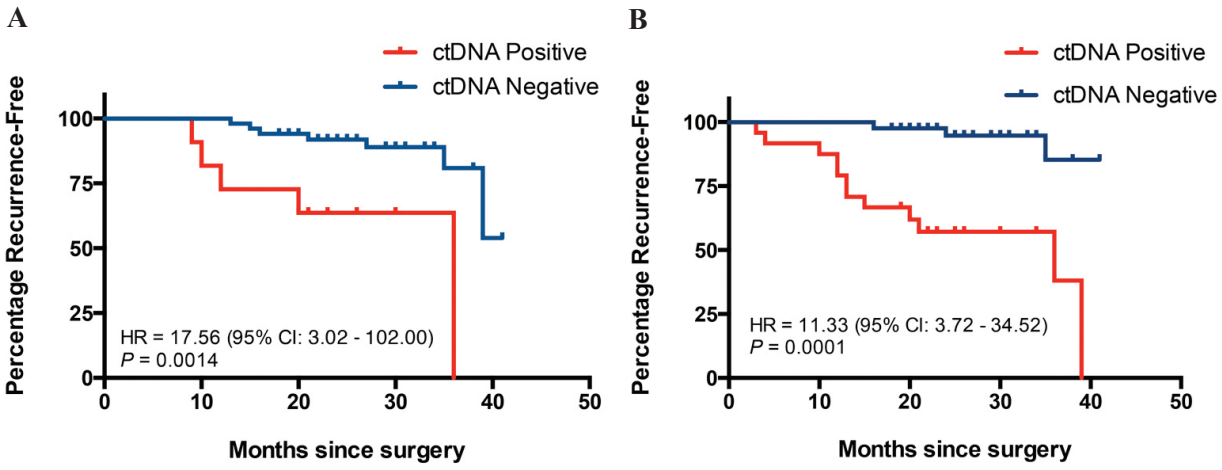


Figure 13. Mutation tracking in serial plasma samples predicts early relapse. (**A**) Disease-free survival according to the detection of ctDNA in the first postsurgical plasma sample [HR, 17.56 (95% CI, 3.02 to 102)]. (B) Disease-free survival according to the detection of ctDNA in serial follow- up samples [HR, 11.3 (95% CI, 3.72 to 34.52)]. P value determined by log- rank test.

Sensitivity in a single postsurgical sample was limited by patients who had undetectable ctDNA in the single postsurgical plasma sample and required serial sampling to detect ctDNA. Of the patients who relapsed in the follow-up period, 50% (7 of 14) had ctDNA detected in the single postsurgical sample and 86.7% (13 of 15) had ctDNA detected by mutation tracking. Of the patients who did not relapse, 88.5% and 61.5% did not have ctDNA detected in either the single postsurgical sample (46 of 52) or by mutation tracking (32 of 52), respectively. Six patients with ctDNA detectable after surgery did not relapse in the follow-up period. Five patients had ctDNA levels that eventually become undetectable during the follow-up. One patient with low risk factors had ctDNA detectable in the first postsurgical sample, with a subsequent increase in ctDNA abundance in serial sampling.

Detection of ctDNA had a median of 11.5 months (range, 3 to 18 months) lead time over radiologic relapse (Figure 14).

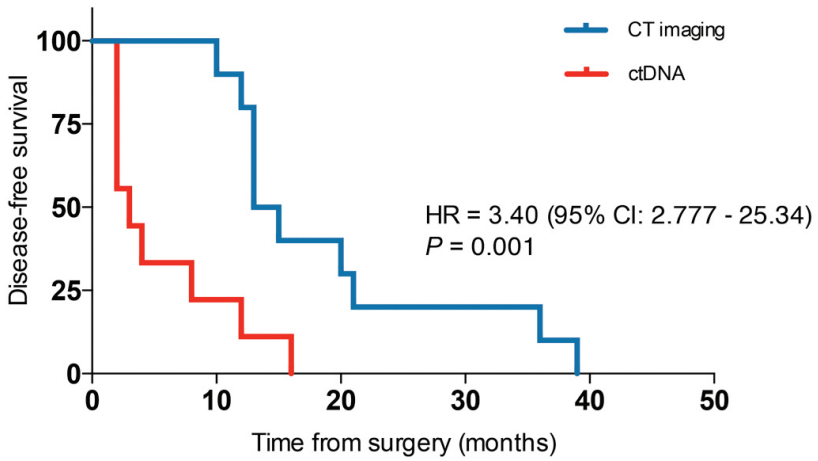


Figure 14. Application of ctDNA analysis for postoperative surveillance. ctDNA in serial plasma after surgery predicted metastatic relapse with a median lead time of 11.5 months over radiologic recurrences.

ctDNA dynamics in patients treated with adjuvant chemotherapy

We then assessed the impact of adjuvant chemotherapy on serial ctDNA status during and after completion of chemotherapy. ctDNA was positive in the postoperative period in 6 of 27 patients with available plasmas after surgery (55 patients received chemo in total), then received at least 3 months of adjuvant chemotherapy. Serial ctDNA concentrations for some of these patients are shown in Figure 15 (A to C). The ctDNA status changed from positive to negative during the initial adjuvant treatment phase in one patient. ctDNA became positive again after completion of chemotherapy in three patients, three of whom later recurred radiologically. In the other two patients, we did not have blood samples collected at this time. Six out of 12 patients who relapsed after treatment harbored ctDNA positive after chemotherapy (three patients did not have a sample taken after chemo and were excluded from this analysis). ctDNA was positive post-chemotherapy in 7 of 25 patients with available plasmas at this point. One of these (stage IIIC) remained recurrence-free at 24 months.

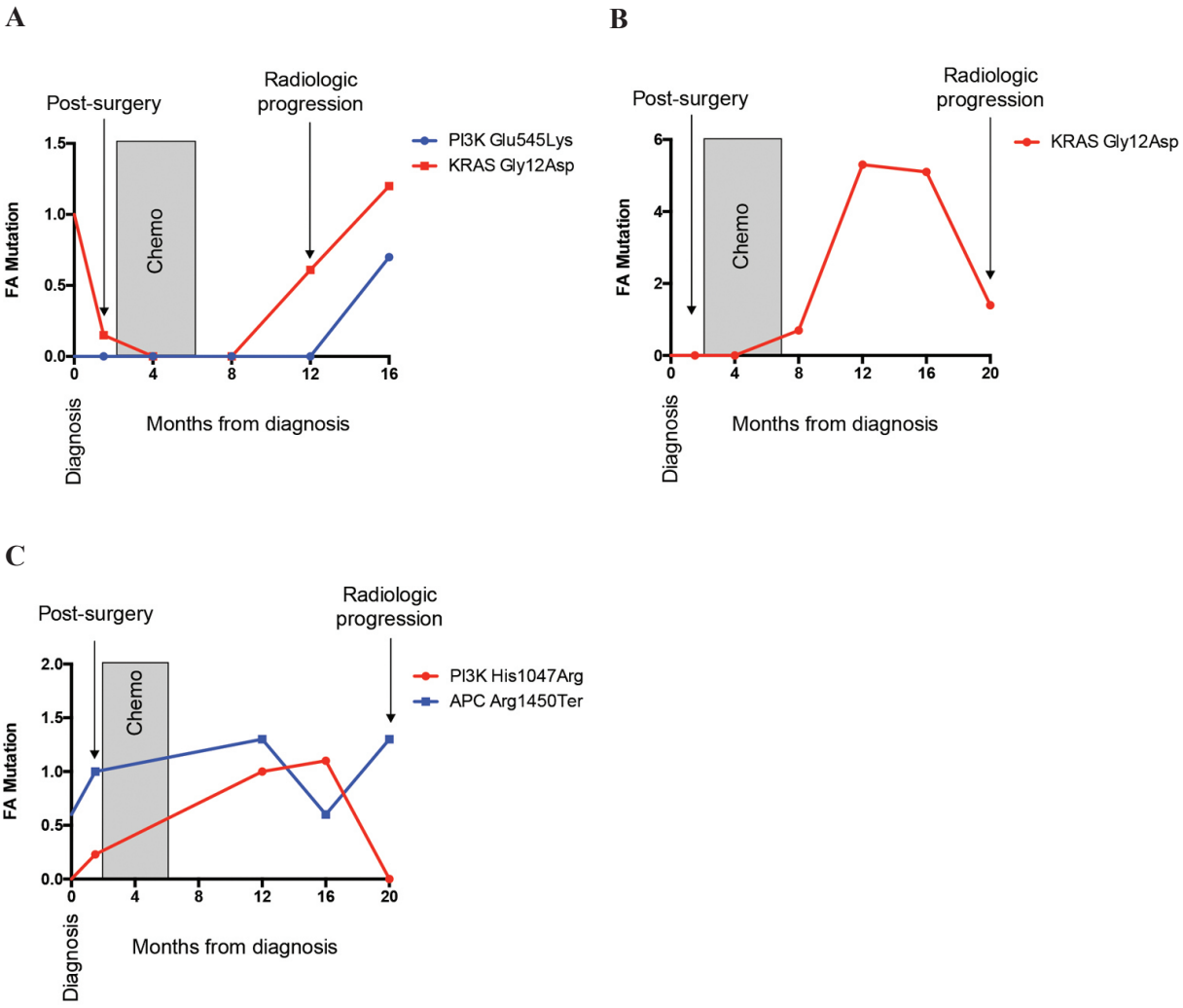


Figure 15. ctDNA status before, during, and after adjuvant chemotherapy. (**A** to **C**) ctDNA concentration (FA: fractional abundance mutation) for three patients with positive postoperative ctDNA who subsequently received adjuvant chemotherapy.

Clinicopathologic variables significantly associated with DFS in univariate analysis were, tumor site, T stage, vascular and perineural invasion. Postoperative ctDNA status had a greater impact on DFS than any individual clinicopathological factors. After multivariable adjustment, postoperative ctDNA status remained an independent predictor of DFS for all patients (HR, 7.02; 95% CI, 1.96 to 25.02) (Table 4).

Variable	Univariate analysis			Multivariate analysis		
	HR	95% CI	*P*	HR	95% CI	*P*
Age, <70 vs ≥ 70	2.081	0.78 – 5.552	0.135			
Sex, male vs female	0.67	0.25 – 1.801	0.427			
Tumor site, right vs left	0.298	0.105 – 0.847	**0.016**			
Tumor differentiation, well-moderate vs poor	0.513	0.117-2.248	0.367			
T stage, T3 vs T4	3.339	1.264 – 8.821	**0.010**	6.84	1.85 – 25.17	**0.004**
Lymph node yield, ≥12 vs <12	2.703	0.357 – 20.474	0.316			
Vascular invasion, no vs yes	0.135	0.048 – 0.383	**0.001**	0.375	0.087 – 1.614	0.188
Perineural invasion, no vs yes	0.380	0.147 – 0.979	**0.038**	1.46	0.247-8.642	0.677
MMR status, deficient vs proficient	2.801	0.779-10.076	0.1			
Stage II risk, high vs low	2.9	1.3 – 5.8	0.218			
Adjuvant chemotherapy, no vs yes	0.408	0.149-1.116	0.072			
Postoperative CEA, normal vs elevated	2.437	0.546-10.867	0.229			
Postoperative ctDNA status, negative vs positive	5.065	1.531-16.756	**0.003**	7.020	1.96 – 25.02	**0.003**

Table 4. Uni- and multivariate analysis. Disease-free survival analysis by clinicopathological variables and postoperative ctDNA status.

Discussion

In this study, we show that ctDNA mutation tracking can detect MRD non- invasively and identify earlier which patients are at risk of cancer recurrence. Serial ctDNA levels during follow-up can precede disease recurrence prior to routine radiographic imaging. In addition, ctDNA measurements had higher sensitivity in detecting recurrence compared with CEA levels. To our knowledge, this is the third study to demonstrate the potential utility of ctDNA in all 3 stages of nonmetastatic CRC [153,154].

Previous studies have shown that ctDNA can be an early marker of disease recurrence in patients with stage II, locally invasive, or metastatic CRC [146,208,209]. The recurrence rate among patients with positive ctDNA post-surgery (11 [58%] of 19) in a previous study of patients locally disease is comparable to findings in the present study (7 [50%] of 14). Our data suggest that the burden of MRD at a single

postsurgical time point is, in some cases, insufficient for its detection in the plasma DNA (50%), or, owing to lack of proliferation and apoptosis in the MRD, there is no release of ctDNA. These findings contrast with a report where ctDNA detection with BEAMing in a single colorectal cancer sample taken after surgical resection of liver metastases offered high predictive potential, potentially because of the high burden of micrometastatic disease in this setting [208]. Nevertheless, we recommend serial sampling to detect the MRD as it proliferates and expands. In this way, detection of ctDNA by mutation tracking increased the sensitivity up to 72.2%.

Also consistent with the previous study is the finding that even patients with positive ctDNA could still be cured by chemotherapy, as demonstrated in two patients. In patients with negative ctDNA immediately post-surgery, the recurrence rate was 9.8% with a follow-up of 27 months in the stage II study, 42.1% with a follow-up of 27 months in the locally disease study and 0% with a follow-up of 49 months in the study for I-III stages in CRC. The recurrence rate in the present study was 50% with a median follow-up of 24.7 months. One possible explanation for the low specificity is the short follow-up, low of stringency of the cutoff for ctDNA positivity used or even the result of some unappreciated technical artifact. However, ctDNA levels of 6 [85.7%] of 7 patients with postoperative ctDNA eventually became undetectable during follow-up. We highlighted the value of serial sampling to increase the specify in this context.

This study also included 29 patients with stage I CRC. None of them relapsed, consistent with the low recurrence rate in this group.

Detection of ctDNA had a median of 11.5 months lead time over radiologic relapse comparable to previous studies. Nevertheless, prospective study is required to determine whether the lead time provided by ctDNA measurements is associated with a better clinical outcome. We used ddPCR to assess plasma samples in this study. This relatively cost-efficient technology represents, along with BEAMing, the most sensitive techniques currently available for detection of known mutations.

It is remarkable to note that sequencing of the plasma DNA taken at relapse is needed to interrogate the genetic characteristics of MRD to identify the lethal clone that may differ in its repertoire of somatic mutations from the dominant clone in the primary cancer. We may detect potentially targetable mutations in the ctDNA that were not in the primary tumor; in some cases, other targetable mutations were lost from primary tumor to ctDNA or resistance mutations during adjuvant treatment.

Under this scenario we are defining a new subgroup of patients with MRD. Novel clinical trials for this subgroup of patients may improve selection for adjuvant therapies and provide new surrogate endpoints for the evaluation of future clinical trial designs in the adjuvant setting [210].

3.3 Evaluation of the prognostic value of a plasma proteomic panel. (Aim 3) Determination of cytokine levels involved in the development of metastases (TGF beta, IL-11 and IL-6) in plasma samples.

To explore if cytokines level could be associated with tumour infiltration, nodal involvement or in general with stage according to TNM system, ELISAs assays were performed. Plasma samples of a prospective cohort of patient were collected at baseline, post-surgery, at follow up and at time of progression as previously discussed. In a first analysis, 132 patients diagnosed with localized CRC and 9 controls were evaluated and studied. ELISAs were carried out as previously described with the aim of analyze three different cytokines, known to have a role in the development and progression of colorectal cancer, such as TGF-beta, IL-6 and IL-11.

The first cytokine analysed was TGF-beta. When tumour infiltration was studied, T parameter according to TNM, it was only possible to detect a slightly differences between T2 and T3 tumours ($p < 0.05$) (Figure 16A). Nevertheless, no statistical differences were observed between N0 versus N positive tumours ($p= 0.93$) (Figure 16B).

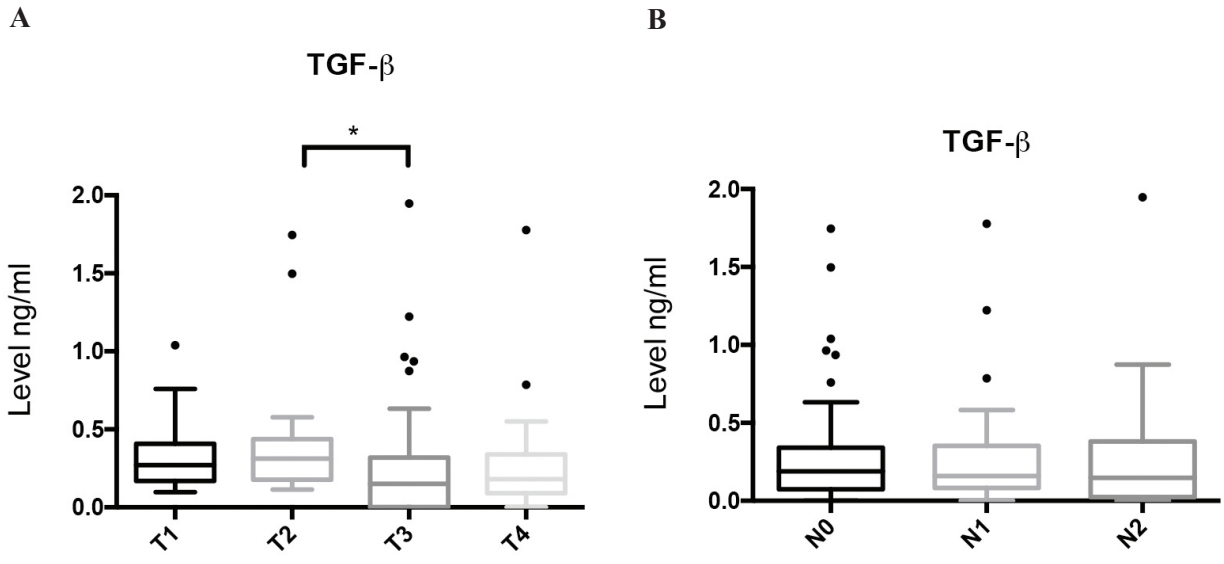


Figure 16. TGF beta plasma levels according to TNM system.

When concentration according to different stages was explored, it was possible to found a statistically significant changes (*Kruskal-Wallis test*) of the concentration between stage I versus stage II or III CRC (p= < 0.05; I and II: p= 0.0012; I and III: p= 0.036) (Figure 17).

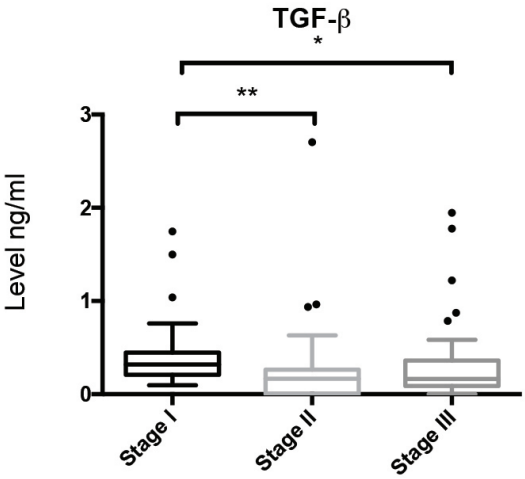


Figure 17. TGF beta plasma levels according to stage.

Likewise, when IL-11 was studied, it was possible to observe that this cytokine variates according to different variables evaluated. Tumour infiltration was analysed and it was possible to see that the level of expression of IL 11 was significantly increased in T3 and T4 patients versus T1 and T2. A statistically significant increment was observed when T2 was analysed versus T4 (p= 0.04) (Figure 18A). When node positive versus negative patients were analysed, a statistically significant difference was also found (p= < 0,05) between N0 and N1. However, these results were not confirmed among N2 population (Figure 18B).

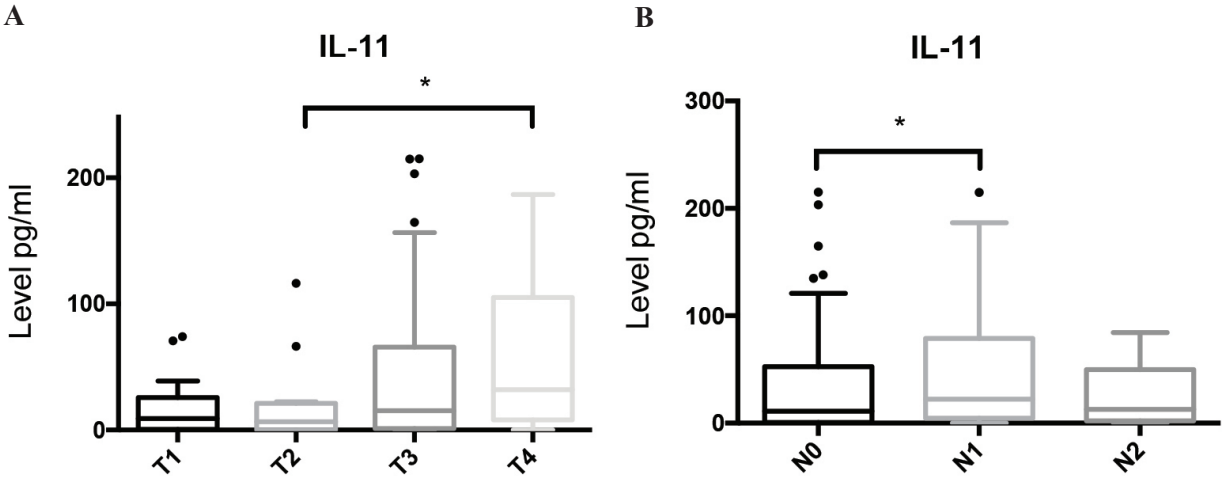


Figure 18. IL11 plasma levels according to TNM system.

Results were also analysed according to the stage and it was possible to observe a statistically significant difference between stage I and III (p= 0.034) (Figure 19).

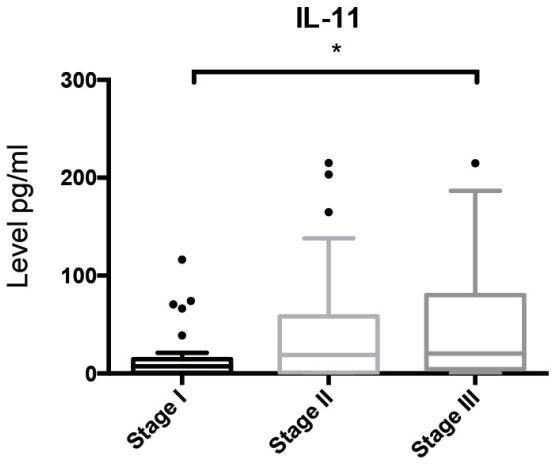


Figure 19. IL11 plasma levels according to stage.

Finally, the last cytokine analysed in our experiment was IL-6. In this case, it was possible to observe that this cytokine increased in patients with T3 and T4 CRC versus T1 (T1 vs T3 p= 0.0026; T1 vs T4 p= 0.0009) (Figure 20A). When the lymph-node involvement was studied, it was also possible to observe a statistically significant difference between N0 and N1 (p= < 0,05), but this result was not confirmed in the N2 population (Figure 20B).

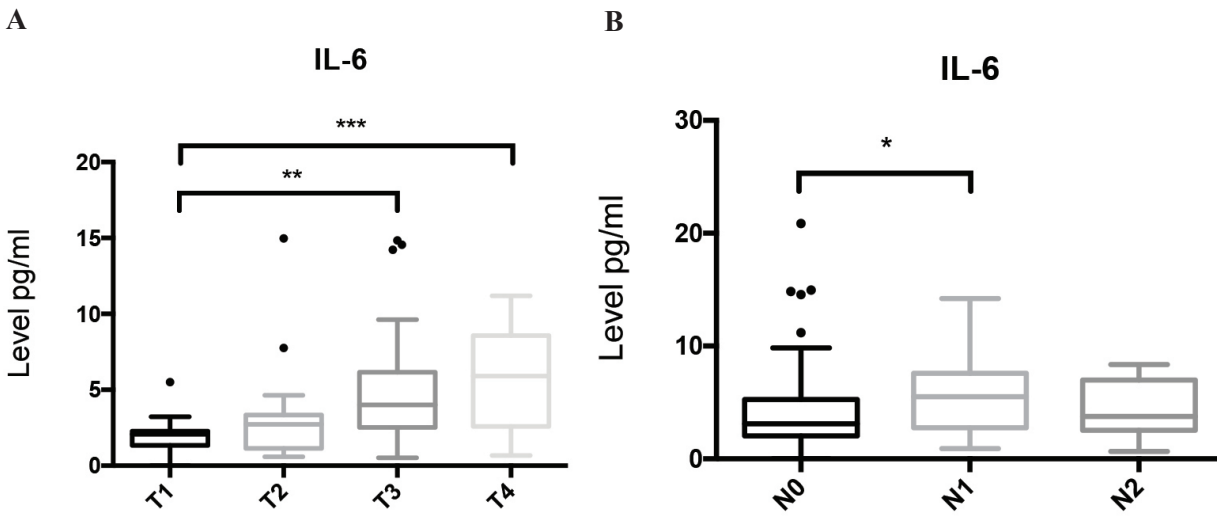


Figure 20. IL6 plasma levels according to TNM system.

Regarding to stage of the disease, it was possible to observe that worse stages according to TNM corresponded to an increase of IL-6 concentration. Specially, the difference between stage I and II was statistically significant (p=0.0056) as the differences between stage I versus stage III (p= 0.0002), with an increasing increment in the concentration of the cytokine, suggesting a lineal regression (Figure 21).

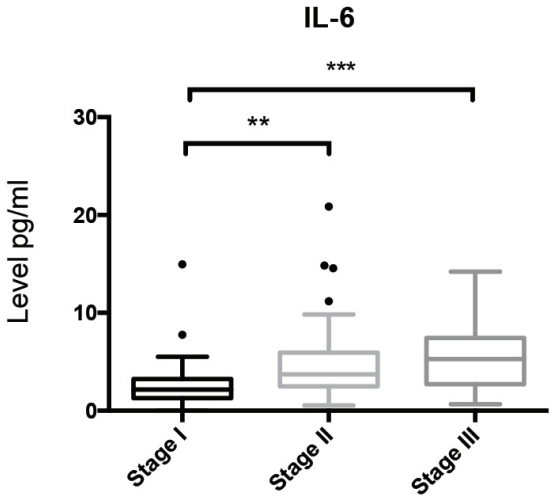


Figure 21. IL6 plasma levels according to stage.

According to the results previously obtained with IL-6 plasma levels, we decided to study a wider population prospectively included into our investigation. TGF beta was no further included into this analysis do the absence of a clear relation within the investigated variables. IL-11 was no further included neither.

In the whole population, when IL-6 concentration was evaluated, it was possible to note some interesting results. In particular, concentration of IL-6 was significantly higher in the plasma of those patients with T3 or T4 tumours versus T1 (T1 vs T3 p= 0.002; T1 vs T4 p= 0.009) (Figure 22A). When nodal involvement was analysed it was also possible to perceive that those patients with node negative disease have lower concentration of IL-6 versus N+ (Figure 22B).

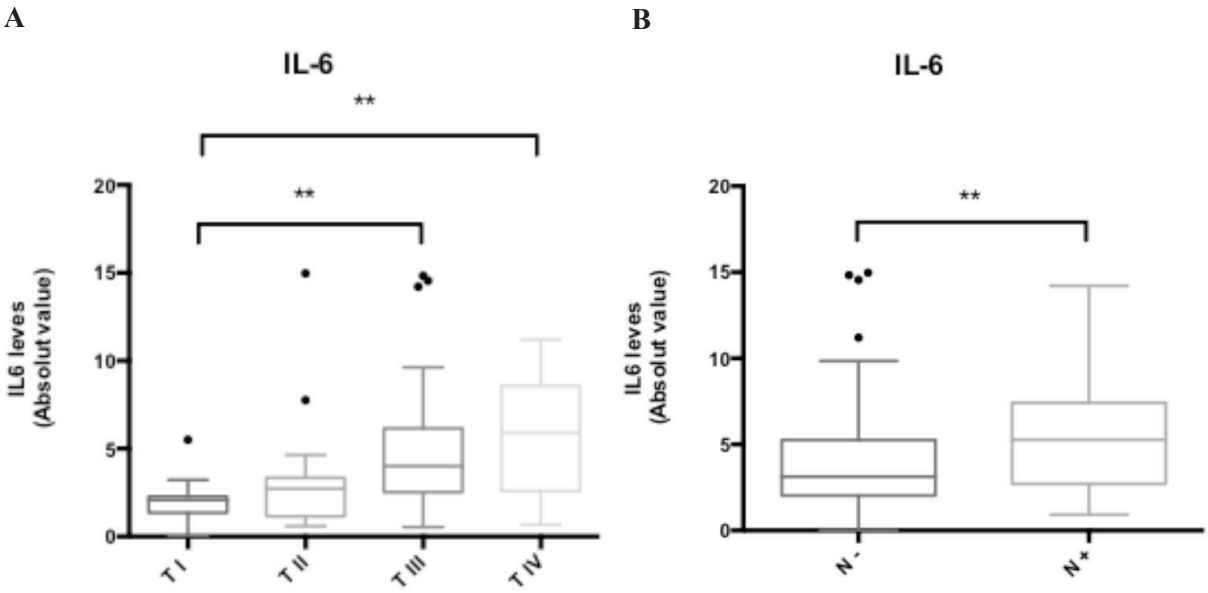


Figure 22. IL6 plasma levels according to TNM system.

When patients were analysed regarding to stage, it was possible to notice a mirror of the previous results. IL-6 was significantly higher in those patients with more advanced disease, particularly the concentration of IL-6 was statistically higher in those patients diagnosed with stage III disease versus those with stage II or I (stage I vs II p= 0.008; stage I vs III p= <0.0001) (Figure 23).

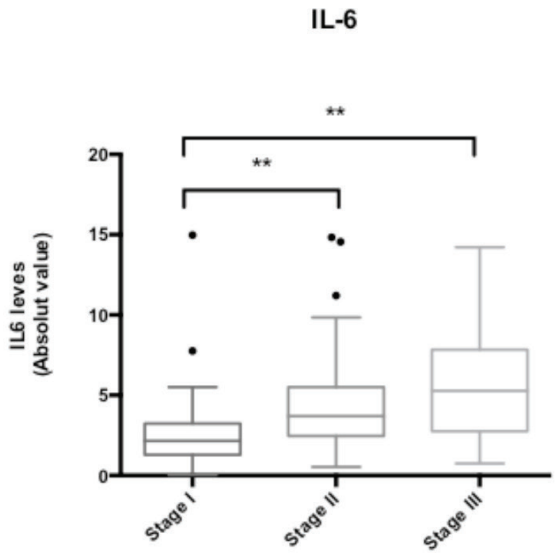


Figure 23. IL6 plasma levels according to stage.

Evaluation of the association between plasma levels of secreted factors with prognosis

A dynamic analysis of cytokine was also performed, evaluating the concentrations at baseline, at follow up and in those patients, who experienced a relapse. Concretely, IL-6 concentrations were evaluated at baseline versus FU and at time of relapse. When concentration of IL-6 was studied in those patients who relapsed it was possible to observe in a statistically significant way that IL-6 concentration at baseline was higher than the level detected at follow up while, at time of relapse, IL-6 concentration increased, suggesting a relation between tumour and IL-6 plasma levels (baseline level vs follow up p= 0.0093; fo versus relapse p= 0,047). Only in one patient this phenomenon was not observed (Figure 24).

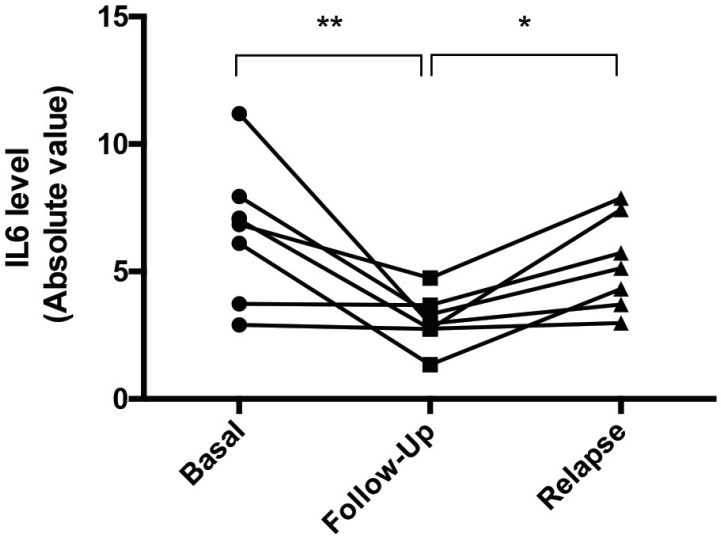


Figure 24. Evolution IL6 plasma concentration during follow up.

Discussion

During recent years, new evidences have highlighted the role of tumor microenvironment as a relevant player in colorectal tumorigenesis and in the development of metastases, being the subject of increasingly intense inquiry.

When peri-tumoral stroma was analyzed, it was found to be different from normal intestinal stroma suggesting that tumor microenvironment could have a relevant role in tumor development and progression. For this reason, evaluating the

role of CAFs and cytokines to better understand the paracrine factors that could be related to cancer development and progression is an interesting field to explore. As reported in several investigations [149,163], the majority of genes able to predict cancer recurrence in CRC patients are expressed not in cancer cells but in CAFs. In this scenario, cytokines are supposed to be important mediators, connecting the inflammatory tumor microenvironment and cancer cell growth. The role of cytokines level in plasma among patients with localized disease and their capability on predict prognosis is a really interesting field of investigation. In our analysis, we selected three different cytokines, TGF-β, IL-11 and IL-6 which were supposed to be related to relapse in CRC patients.

TGF beta role in CRC has been the objectives of relevant investigation which have finally underlined the relevance of this protein and the implication in CRC development and progression through the activation of a specific signaling pathway and modification of microenvironment. In our experiment cytokine levels were evaluated with ELISA assays. TGF beta plasma levels detected in plasma were not related with T and N parameters. No relation with relapse was observed. When the activated isoform was evaluated the results obtained were no longer different from the previous one, suggesting that in our series, in case of relapse TGF beta was not increased. These results are apparently in contrast with the others already published in the previous literature where several patients diagnosed with CRC cancer and other solid tumors were analyzed an increase of TGF beta was detected in case of progression disease [211]. The most relevant difference between these studies and the one that we propose is the stage of patients so that it would have been possible that the huge of increase of TGF beta is no so relevant in localized disease. Moreover, it would be possible that the lack of detection of increased level could be caused by the absence of intracellular conversion in the activated isoform. We hypothesize that TGF beta probably is activated within tumor cell instead of plasma.

The second cytokines that we studied was IL-11. This cytokine has several functions and is also supposed to be activated by TGF beta signaling. IL-11 was

described to reduce cancer cell clearance during initial colonization and augmented the metastatic potential of CRC cells. Accumulating evidence alludes to a role for IL-11 signaling in tumor development although the mechanisms underlying IL-11 biology in CRC remain largely unknown. In our experiments, it was no longer possible detect the relationship between its plasma level and relapse or clinical parameters. These results are in line with the one obtained with TGF beta, further investigation to clarify this point are needed.

Finally, when the ELISA assay was used to evaluate the level of IL-6, interestingly it was possible to note a correlation with T and N parameters (T1 vs T3 p= 0.002; T1 vs T4 p= 0.009; N0 vs N+ p= 0.005) and with stage of the disease. IL-6 was statistically higher in those patients diagnosed with stage III disease than those with stage II or I (stage I vs II p= 0.008; stage I vs III p= <0.0001). These results are of interest due to the fact that IL-6 seems to directly relate to poor prognosis parameters such as T increase and nodal involvement. Nevertheless, the most important result was its correlation with relapse. When concentration of IL-6 was studied in those patients who relapsed it was possible to perceive in a statistically significant way that at baseline IL-6 concentration was higher versus the level detected at follow up while, at time of relapse, IL-6 concentration increased, suggesting a relation between tumour and IL-6 plasma level (p= 0.0093 and p= 0.047, respectively). The value of IL-6 to be considered as positive or negative derived form an AUC of the ROC curve.

IL-6 is an important pro inflammatory cytokine, IL-6 signal transducer (also known as GP130) and STAT3 signaling in tumor epithelial cells promoted the survival of CRC cells during colonization and was required for efficient metastasis initiation. Moreover, when patients who experienced an increase of IL-6 were evaluated according to CMS subgroup, we observed that patients presenting IL-6 high level were the ones belonging to CMS1 and CMS4 (42,2% and 57,8%). These data confirm the role in immunomodulation of this cytokine in these two different

subtypes of patients where probably the microenvironment play a principal role in tumor development and progression.

In our study, although the multivariate Cox analysis only suggest a trend towards the statistically significance presumably due to insufficient data [Table 5], further investigation is needed to clarify the role of IL-6 in these patients as biomarker of relapse and probably as a novel target for therapeutic approaches.

3.4 Classification of patients according to molecular subtypes (CRCAssigner and Consensus Molecular Subtypes (CMS) using a customised 38-gene assay for nCounter platform (NanoString Technology). (Aim 4)

To evaluate whether gene expression signatures add clinically relevant information to existing clinical subgroups in early stage we used a customized 38-gene assay that demonstrated the validation of a custom NanoString nCounter platform-based biomarker assay (NanoCRCA) to stratify CRC into subtypes [212].

117 out of 150 patients from our cohort of primary tumour samples were collected prior to treatment. All the CRCA subtypes were present in this cohort, and all samples were successfully classified, with none showing undetermined subtype in the nearestCRCA group (Figure 25). We observed a non-uniform distribution of the subtypes with enterocyte contributing 25% of all the samples followed by globet-like (24%), TA (18%) and inflammatory (18%) subtypes. Stem-like (15%) subtype was lower in frequency in this cohort of samples.

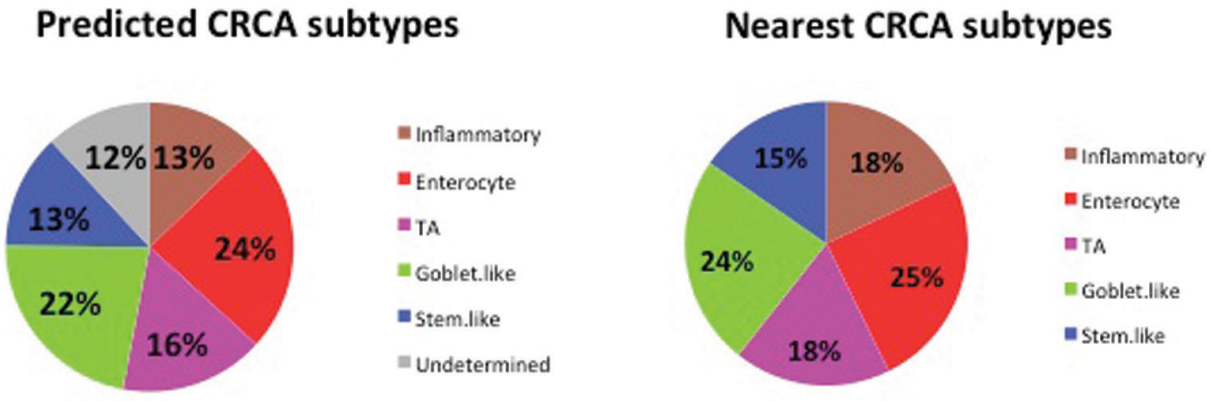


Figure 25. NanoCRCA assay: pie charts showing the proportion of different subtypes (including undetermined samples).

We additionally compared the NanoCRCA assay with the microarray-based CMS subtypes in this cohort. We classified our cohort into CMS subtypes using the published CMS classifier [79]. Two different algorithms for CMS classification available online in the R¨CMSclassifier¨ package were used: one suitable for population-based studies (random forest classifier) and one suitable for single-sample

prediction (Pearson correlation-based classifier, SSP). We successfully classified the samples into all of the CMS subtypes: 37.0% were CMS4 (generally associated with stem-like); 28% of CMS2 (enterocyte and TA); 20% of CMS1 (inflammatory); and 15% of CMS3 (goblet-like). However, we found 32% samples with undetermined subtype in the predicted classifier (Figure 26; defined as "mixed" samples in the original publication [79]).

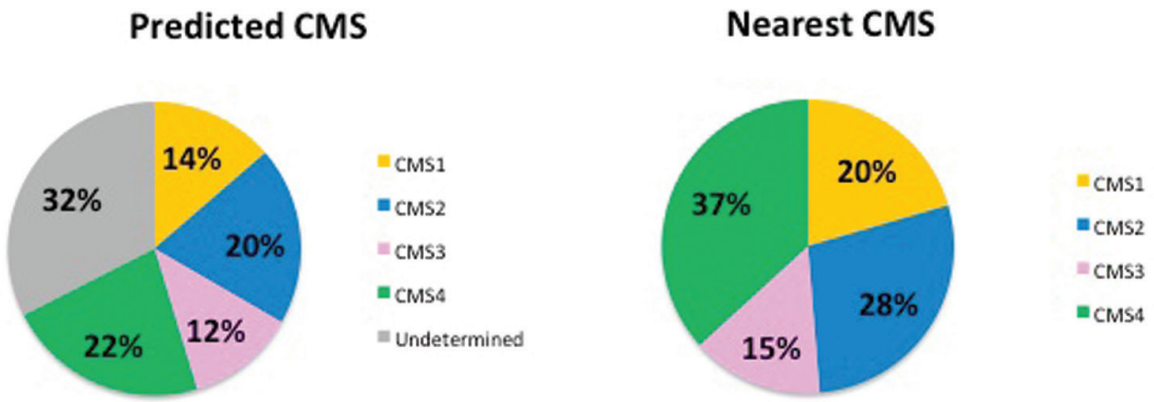


Figure 26. NanoCRCA assay, its comparison with the CMS classifier: pie charts showing the proportion of different subtypes (including undetermined samples) from microarray-based CMS classifications.

The standard and modified protocols used in this cohort showed high Pearson correlation (>0.88) for gene expression [212]. NanoCRCA classified FFPE samples into all five CRCA. We demonstrated high concordance across NanoCRCA assay and with CMS subtypes.

In a univariate analysis, CMS subtypes were significantly associated with CDX2 expression (loss in CMS1, p=0.03), IL-6 levels (high in CMS1 and CMS4, p=0.002) and perineural invasion (present in CMS1 and CMS4, p=0.001).

After a median follow-up of 24.7 months, we showed CMS1 and CMS4 subtypes were significantly associated with relapse (p=0.016) Figure 27.

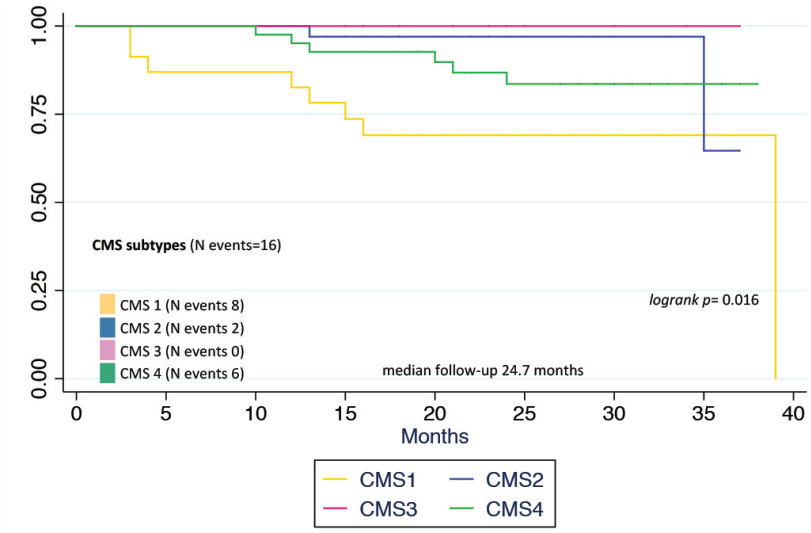


Figure 27. Kaplan-Meier survival analyses for disease-free survival according to nearest CMS subtype.

3.4.3 Correlation of the molecular subtypes with ctDNA in pations with localized CRC

When we analyze the value of integrating postoperative ctDNA status into CMS classification, we note that the immediate detection of ctDNA status after surgery was related to poor prognosis in the CMS4 subtype (p<0.0001) Figure 28.

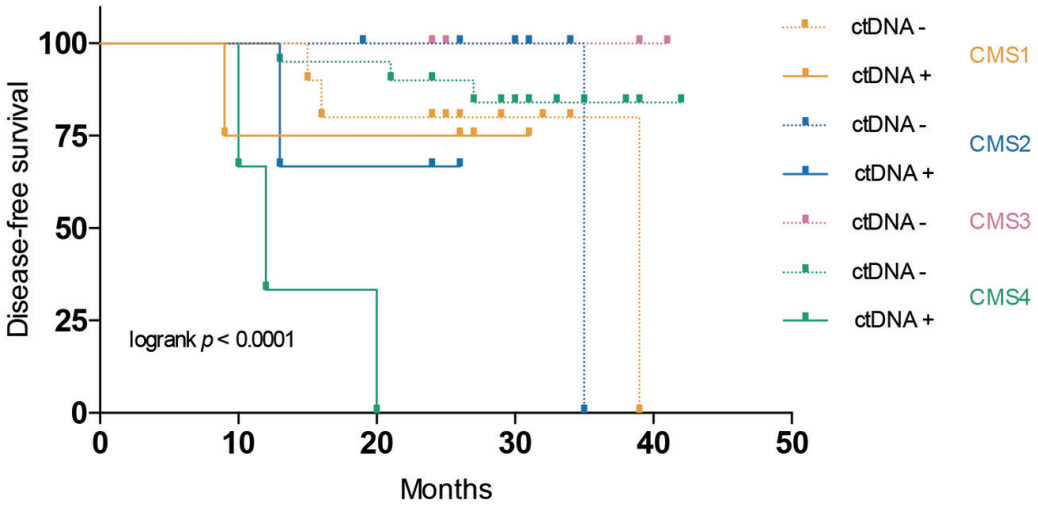


Figure 28. Kaplan-Meier survival analyses for disease-free survival according to ctDNA post-surgery status and CMS subtypes.

Additionally, we showed that the detection of postoperative ctDNA in serial samples (called mutation tracking) was statistically significant associated with poor prognosis in the CMS1 and CMS4 subtypes (p=0.0012) Figure 29.

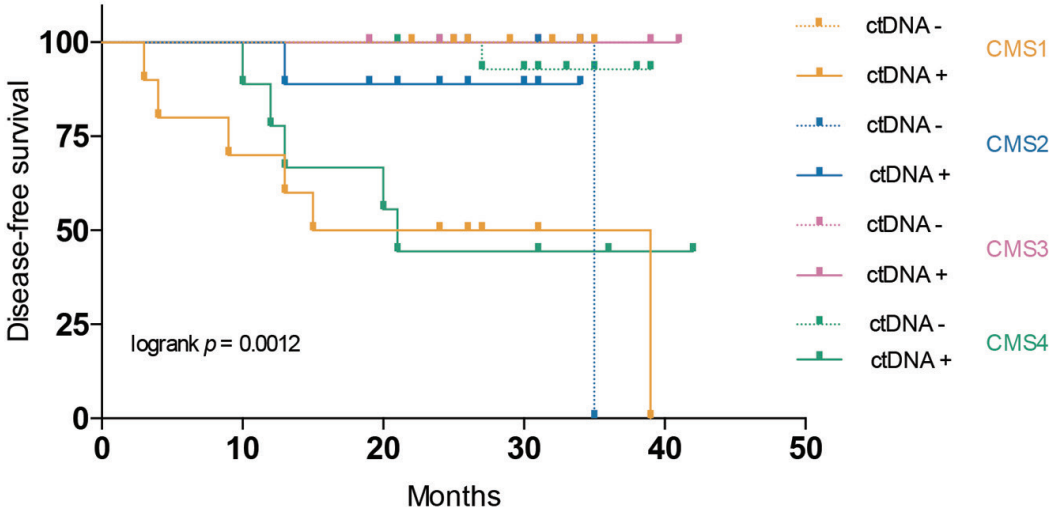


Figure 29. Kaplan-Meier survival analyses for disease-free survival according to ctDNA mutation tracking status and CMS subtypes.

Discussion

CRC is a highly heterogeneous disease. Prognosis has been based essentially on tumor extension categorized by the TNM staging method. Proposed cancer therapies are based on these classifications including conventional chemotherapies and personalized therapies regarding mutated genes. Although different multigene assays as Oncotype DX, ColoPrint and ColDX have demonstrated independent prognostic value in localized CRC, their use is currently not recommended due to unclear clinical utility over current risk stratification factors and due to the lack of value in predicting treatment benefit. Whether gene expression signatures add clinically noted information to existing clinical subgroups in early or advanced stages remains controversial.

As a result of CMS classification for 117 cases, 24 (20%), 33 (28%), 17 (15%), and 43 (37%) were classified into CMS1, CMS2, CMS3, and CMS4, respectively in our cohort. This stratification is in line with previous studies.

In our prospective cohort, the microsatellite instable-enriched subgroup (CMS1) and the mesenchymal subgroup (CMS4) were statistically associated with shorter DFS (Figure 27). These findings differ from previous data where CMS1 subtype was related to good prognosis. One explanation could be that the lack of CDX2 expression were mostly observed in the CMS1 group in our cohort. In contrast, the canonical (CMS2) and metabolic (CMS3) subtypes with intermediate prognosis exhibited better DFS as expected. Recently, the CMS subtypes were evaluated as independent prognostic factors of survival demonstrating consistent results in correlative studies of phase III clinical trials in line with our prospective study [213,214]; however, conflicting results were shown when tested as predictive factors of benefit from standard therapy in the metastatic setting.

We must take into account that the CMS subtypes were developed in the context of primary CRC and its applicability to metastatic setting has not yet been fully established. Moreover, owing to intratumor heterogeneity same authors are questioning the robustness of CMS subtypes in tissue, demonstrating a high proportion of unclassified samples [215] being in our cohort of 32% in the predicted algorithms probably due to a heterogeneous disease with more than one subtype present in the same sample. In the line with previous studies, using the ´nearestCMS´ subtype classification of the previously undetermined 44.7% and 26.5% were classified as CMS4 and CMS2 (the most heterogeneous group), respectively. Therefore, the choice of CMS predictor algorithm is crucial in certain contexts [216].

In the era of targeted therapies, particularly immunotherapies that are dependent on the composition of the TME, we undertook to integrate molecular and immune features of CRCs. We observed that the higher levels of IL6 at baseline were in the CMS1 and CMS 4 subtypes (p=0.002). Probably, this fact is due to MSI colorectal cancers are more immunogenetic than MSS tumors and the CMS4 subtype is enriched for mesenchymal and stromal gene signatures, which may be derived from surrounding cells rather than being cancer specific.

On the other hand, CMS subtypes were significantly associated with CDX2 expression being lost in CMS1 (p=0.03). It is remarkable to note that 5 (62.5%) of 8 relapsed patients classified as CMS1 had lost CDX2 expression. In line with Pilati C data [217] our study showed that lack of CDX2 expression is only present in the MSI-immune tumors (CMS1) and the mesenchymal subgroup (CMS4) and not in CMS2 and CMS3 colon cancer. However, contrary to our data, which demonstrated that loss of CDX2 expression, is statistically associated with worse prognosis not only in the CMS4 group but also in CMS1 patients.

To our knowledge this is the first report to correlate the molecular subtypes with ctDNA status in patients with localized CRC. We demonstrated that ctDNA positive in serial plasma samples during the follow-up was statistically significant associated with shorter DFS in the CMS1 and CMS4 subtypes (p=0.0012). Again, we defined two subgroups with poor prognosis and maybe we should take into account the integration of ctDNA status into trancriptomic analysis in future studies.

We highlight that further analysis is warranted to put all this data in the right context and to take a further step towards precision medicine for patients diagnosed with localized CRC.

3.5 Integration of a comprehensive molecular assessment to elaborate a valuable prognostic/predictive multi-omic model. (Aim 5)

As we previously mentioned, the acquisition of cancer hallmarks requires molecular alterations at multiple levels including genome, epigenome, transcriptome, proteome, and metabolome.

Based on data before published, we hypothesized that the integration of genomic, transcriptomic and proteomic analysis into one could enhance the sensitivity to detect MRD and predict recurrence in patients diagnosed with localized CRC and so stratify them in subtypes according to multi-omics features. This could help to detect patients at high risk of relapse and resistance mechanisms to standard treatment with the ultimate goal of personalizing therapy in relapsed patients.

To try answering this point, we integrated all previous reported data into statistics analysis. Firstly, we performed an univariate analysis on data that we collected such as age, sex, tumor site, tumor differentiation, T stage, lymph node yield, vascular and perineural invasion, MMR status, stage II risk, adjuvant chemotherapy, postoperative CEA, postoperative ctDNA, postoperative mutation tracking, CDX2 status, IL6 level, CMS classification and *KRAS*, *PIK3CA*, *BRAF* and *CTNNB1* mutation status. The variables which showed statistically significance related to shorter DFS were right sided site (p=0.016), T4 stage (p=0.010), positive lymph nodes (p=0.003), vascular invasion (p=0.001), perineural invasion (p=0.038), positive postoperative ctDNA (p=0.003), positive mutation tracking (p=0.003), lack of CDX2 expression (p<0.0001), high IL6 levels at baseline (p=0.038) and CMS1 and CMS4 subtypes (p=0.029). Immediately, we performed a multivariate analysis finding shorter DFS statistically significance in patients with T4 stage (p=0.041), positive postoperative ctDNA (p=0.006) and/or positive mutation tracking (p=0.004) and lack of CDX2 expression (p=0.04).

Overall, 18 (12%) of the 150 patients relapsed. Eight (44.4%) of those 18 patients showed lack of CDX2 expression. Three (30%) of the 10 patients with expressed CDX2 were classified as CMS1 subtype, five (50%) as CMS4 subtype and the other two patients had not CMS status available.

We found that CDX2 negative represented 5.3% (8/150) patients of our study population. All patients were diagnosed with stage III CRC cancer and the majority of the tumors (75%) with CDX2 loss were located in right side of the colon. Although, 87.5% of patients had received adjuvant therapy (only 28.6% patients received oxaliplatin based therapy), 6 (75%) of the 8 patients relapsed. CDX2 negative samples were enriched in CMS1 (83.3%) and CMS4 (16.7%) groups, *BRAF* (42.9%) and *KRAS* (42.9%) mutations and vascular/perineural invasion (75%). Among the 150 tumors, we determined CMS classification for 117 samples. Interestingly, 6 (100%) of the 6 CDX2 loss patients with available CMS group belonged to the CMS1 and CMS4 groups. These two groups are related to MSI and immune infiltration (CMS1) and EMT phenotype (CMS4).

Globally CDX2 expression was an independent prognostic factor for DFS (p=0.04) in multivariate analysis adjusted for TNM stage, vascular/perineural invasion, IL6 levels at baseline, tumor location and postoperative ctDNA status, as previously reported (Table 5).

Variable	Univariate analysis			Multivariate analysis		
	HR	95% CI	*P*	HR	95% CI	*P*
Age, <70 vs ≥ 70	2.081	0.78 – 5.552	0.135			
Sex, male vs female	0.67	0.25 – 1.801	0.427			
Tumor site, right vs left	0.298	0.105 – 0.847	**0.016**	0.315	0.048-2.055	0.227
Tumor differentiation, well-moderate vs poor	0.513	0.117-2.248	0.367			
T stage, T3 vs T4	3.339	1.264 – 8.821	**0.010**	7.316	1.08 - 49.54	**0.041**
Lymph node Positive vs negative	4.706	1.514-14.631	**0.003**	5.635	0.379-83.84	0.209
Lymph node yield, ≥12 vs <12	2.703	0.357 – 20.474	0.316			
Vascular invasion, no vs yes	0.135	0.048 – 0.383	**0.001**			
Perineural invasion, no vs yes	0.380	0.147 – 0.979	**0.038**	0.111	0.008-1.56	0.103
MMR status, deficient vs proficient	2.801	0.779-10.076	0.1			
Stage II risk, high vs low	2.9	1.3 – 5.8	0.218			
Adjuvant chemotherapy, no vs yes	0.408	0.149-1.116	0.072			
Postoperative CEA, normal vs elevated	2.437	0.546-10.867	0.229			
Postoperative ctDNA status, negative vs positive	5.065	1.531-16.756	**0.003**	17.105	2.247-130.18	**0.006**
CDX2 status present vs low	12.682	4.637 – 34.686	**<0.0001**	69.1	3.869-1234.4	**0.004**
Il6 level, high vs low	3.464	0.994 – 12.077	**0.038**	3.689	0.396-34.322	0.251
CMS classification CMS1 vs CMS2-3-4	2.756	0.918- 8.629	**0.016**	1.487	0.861	0.155
KRAS status, wild type vs mut	0.512	0.195 – 1.345	0.166			
PIK3CA status, wild-type vs mut	0.529	0.195-1.435	0.203			
BRAF status, wild type vs mut	0.538	0.120-2.403	0.409			
CTNNB1 status, wild type vs mut	0.460	0.059–3.588	0.448			

Table 5. Uni- and multivariate analysis. Disease-free survival analysis by clinicopathological variables and postoperative ctDNA status.

We thus examined the prognostic impact of CDX2 separately in CMS1 and 4 groups and found that CDX2 status was highly prognostic in CMS1 (p=0.0001) and also CMS4 patients for relapse (Figure 30).

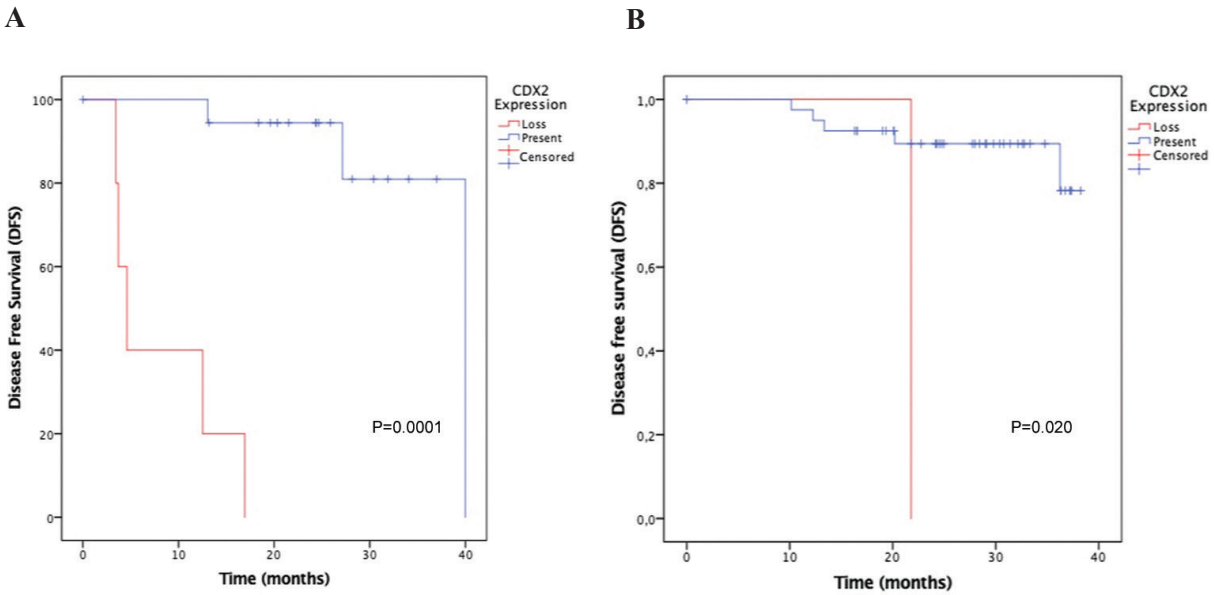


Figure 30. Kaplan-Meier survival analyses for disease-free survival according to CDX2 expression status and CMS subtypes.

We did not find MSI phenotype overrepresented in CDX2 negative group (3/8 patients). We did not discover a role of KRAS/BRAF/PIK3CA mutations in localized CRC (p >0.05).

We are currently working on combining the detection of ctDNA and proteins in plasma samples to enhance the likelihood of identifying patients at very high risk of relapse.

Discussion

We wanted to examine whether the sensibility to identify patients at high risk of recurrence could be increased adding another variable to postoperative ctDNA status. Here, we combined protein biomarkers with genetic biomarkers to increase sensitivity without substantially decreasing specificity. Other cancer biomarkers, such as metabolites, mRNA transcripts, miRNAs, or methylated DNA sequences could be similarly combined to increase sensitivity and detection of MRD. Such multianalyte tests would not be not meant to replace other non-blood-based screening tests, such as those for breast or colorectal cancers, but to provide additional information that

could help identify those patients most likely to harbor a malignancy. The study was limited by the sample size, involving only blood samples from 85 patients. We are presently working on the statistical analysis of this point.

Here we identified a loss of CDX2 expression in 2 CMS subgroups: CMS1, characterized by MSI, immune infiltration and good prognosis; and CMS4, showing a mesenchymal/stem cell phenotype and a poor prognosis. Both these groups have been associated with activation of the TGF beta pathway, particularly strong in CMS4 [159].

In our study we showed that, CDX2 was an independent prognostic factor, loss of CDX2 expression allowed to identify a group of patients with a particularly poor prognosis (CMS1 y 4 subtypes). The lack of CDX2 expression was a poor prognostic factor regardless of *BRAF* mutational status or MMR status.

Somatic mutations in the PI3K/AKT1/MTOR pathway play a vital role in carcinogenesis. Approximately 40%, 8%, and 15% of CRCs harbor activating mutations in *KRAS*, *BRAF*, and *PIK3CA* respectively, making it the most frequently mutated genes in CRC. However, our current study demonstrated that the *KRAS*, *BRAF* and *PIK3CA* mutation status was not significantly associated with CRC patient survival in the localized setting. These findings are in line with the literature [218]. Tumor behavior may depend on multiple differing driver events as well as interaction of many molecular alterations in tumor, not solely on *KRAS*, *BRAF*, and *PIK3CA* mutation status. Moreover, host factors such as immune response to tumor may influence prognosis. Thus, it is not surprising to find that any driver mutation (such as *PIK3CA* mutation) has no substantial prognostic role. Previous studied have outlined the role of *KRAS*/*BRAF*/*PIK3CA* mutation status for predicting outcome in mCRC [219].

Here, we showed that postoperative ctDNA status can detect MRD non-invasively and identify earlier which patients are at risk of cancer recurrence. Additionally, the detection of ctDNA by mutation tracking increased the sensitivity up to 72.2%. Detection of postoperative ctDNA or by mutation tracking was a significant predictor of recurrence in a multivariable model [p=0.006 and p=0.004, respectively]. To our knowledge this is the first study to consider the role of mutation tracking in CRC.

In conclusion, combination of postoperative ctDNA status, levels of IL6 in plasma, the CMS classification and lack of CDX2 expression could be a useful marker to identify high-risk patients with poor prognosis. CDX2 does not seem to bear the same prognostic information in the different CMS subgroups and could represent synergistic biomarkers. However, CDX2 was an independent prognostic factor in the multivariate analysis to be used in combination. Whether patients with CMS1/4-CDX2 negative colon cancer might be treated differently than others due to their poor outcome require further investigation. Our data suggest that CMS and CDX2 may be an important stratification factor for future adjuvant trials.

CONCLUSIONS

4. Conclusions

1. Our custom-designed targeted next generation sequencing panel detects mutations in colon cancer samples with high accuracy and sensitivity.

2. Detection of plasma ctDNA after surgical resection with curative intent implies the presence of minimal residual disease and higher risk of recurrence. Moreover, plasma mutation tracking can properly identify those patients at higher risk of relapse with a median of 11.5 months lead time over radiologic identification.

3. High levels of interleukin 6 in plasma at baseline are associated with shorter disease-free survival in localized colon cancer.

4. Consensus Molecular Subtypes 1 and 4 are associated with shorter disease-free survival in localized colon cancer.

5. In a mul-tiomics approach, mutation tracking in serial plasma samples detects minimal residual disease in patients with localized colon cancer and identifies those at high risk of recurrence regardless of stage, lack of CDX2 expression and CMS subtype.

SUMMARY

5. Summary

El cáncer colorrectal (CCR) es un importante problema de salud pública. Más de 1.2 millones de pacientes se diagnostican de CCR cada año, y más de 600.000 mueren a causa de la enfermedad. El pronóstico de los pacientes con CCR ha mejorado durante las últimas décadas, con una supervivencia a los 5 años que ha alcanzado casi el 65% en países desarrollados. El estadio al diagnóstico es actualmente el factor pronóstico más importante. Aunque la clasificación clásica de estadiaje proporciona información pronóstica y guía las decisiones terapéuticas, no disponemos hoy en día de marcadores predictivos de la respuesta y del resultado clínico de la terapia individual del paciente. Esto es relevante en el paciente con CCR localizado - estadio II y estadio III - en los que el tratamiento estándar es la cirugía seguida de quimioterapia adyuvante en estadio III y en estadio II con factores de riesgo adicionales. Sin embargo, una proporción significativa de casos no se beneficia del tratamiento adyuvante, con un 20-50% de recaídas a pesar de la terapia inicial. La identificación de marcadores predictivos y pronósticos más allá del estadiaje es crucial para ayudar a seleccionar mejor los pacientes con alto riesgo de recaída y establecer potenciales estrategias terapéuticas para optimizar el tratamiento adyuvante.

La biopsia líquida es la determinación de las características moleculares en una extracción de sangre del paciente, incluyendo el análisis de DNA tumoral circulante (ctDNA). La biopsia líquida proporciona de forma fácil y poco invasiva, información relevante sobre (1) la presencia de tumor en el organismo incluso cuando es indetectable macroscópicamente mediante técnicas de imagen convencionales (2) mecanismos moleculares de resistencia al tratamiento adyuvante. En base a esto, el objetivo del presente proyecto es determinar el valor pronóstico y predictivo de la biopsia líquida en pacientes con CCR localizado y su integración dentro de una estrategia multiómica a través del análisis secuencial de tejido y biopsias líquidas. Para ello, analizaremos el impacto clínico de la biopsia líquida basal en 150

pacientes con CCR localizado, así como del seguimiento con análisis seriados de biopsia líquida, e integraremos dicha información dentro de un análisis multiómico con la última finalidad de mejorar el pronóstico y como primer paso para optimizar los tratamientos de precisión en ensayos clínicos de tratamiento adyuvante en CCR.

REFERENCES

References

1. Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J et al. Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. Eur J Cancer 2013;49:1374-403.
2. Ferlay J, Soerjomataram I, Dikshit R et al. Cancer indicence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer 2015,136:E359-E386.
3. Center MM, Jemal A, Smith RA et al. Worldwide variations in colorectal cancer. CA Cancer J Clin 2009;59:366-78.
4. Bray F. Transitions in human development and the global cancer burden. In: Steward BW, Wild CP, eds. World Cancer Report 2014. Lyon: International Agency for Research on Cancer 2014,54-68.
5. Center MM, Jemal A, Ward E. International trends in colorectal cancer incidence rates. Cancer Epidemiol Biomarkers Prev 2009;18:1688-94.
6. Murphy CC, Harlan LC, Lund JL et al. Patterns of colorectal cancer care in the United States: 1990-2010. J Natl Cancer Inst 2015;107.
7. Vasen HF, Tomlinson I, Castells A. Clinical management of hereditary colorectal cancer syndromes. Nat Rev Gastroenterol Hepatol 2015; 12: 88-97.
8. Tiwari AK, Roy HK, Lynch HT. Lynch syndrome in the 21st century: clinical perspectives. QJM 2016;109:151-158.
9. Van Lier MG, Leenen CH, Wagner A et al. Yield of routine molecular analyses in colorectal cancer patients ≤70 years to detect underlying Lynch syndrome. J Pathol 2012;226:764-74.
10. Jess T, Rungoe C, Peyrin-Biroulet L. Risk of colorectal cancer in patients with ulcerative colitis: a meta-analysis of population-based cohort studies. Clin Gastroenterol Hepatol 2012;10:639-45.
11. Castaño-Milla C, Chaparro M, Gisbert JP. Systematic review with meta-analysis: the declining risk of colorectal cancer in ulcerative colitis. Aliment Pharmacol Ther 2014;39: 645-59.
12. World Cancer Research Fund/American Institute for Cancer Research. Continuous update project report. Food, nutrition, physical activity, and the prevention of colorectal cancer. Diet and Cancer Report [online], http://www.dietandcancerreport.org/cancer_resource_center/downloads/cu/Colorectal- Cancer-2011-Report.pdf (2011).
13. Guraya SY. Association of type 2 diabetes mellitus and the risk of colorectal cancer: a meta-analysis and systematic review. World J Gastroenterol 2015;21:6026-31.
14. Fedirko V, Tramacere I, Bagnardi V et al. Alcohol drinking and colorectal cancer risk: an overall and dose-response meta- analysis of published studies. Ann Oncol 2011;22: 1958-72.
15. Liang PS, Chen TY, Giovannucci E. Cigarette smoking and colorectal cancer incidence and mortality: systematic review and meta-analysis. Int J Cancer 2009;124:2406-15.
16. Song M, Garrett WS, Chan AT. Nutrients, foods, and colorectal cancer prevention. Gastroenterology 2015;148:1244-60.
17. Arem H, Moore SC, Park Y et al. Physical activity and cancer-specific mortality in the NIH-AARP Diet and Health Study cohort. Int J Cancer 2014;135:423-31.
18. Algra AM, Rothwell PM. Effects of regular aspirin on long-term cancer incidence and metastasis: a systematic comparison of evidence from observational studies versus randomised trials. Lancet Oncol 2012;13:518-27.
19. Doubeni CA, Major JM, Laiyemo AO et al. Contribution of behavioral risk factors and obesity to socioeconomic differences in colorectal cancer incidence. J Natl Cancer Inst 2012;104:1353-62.
20. Erdrich J, Zhang X, Giovannucci E et al. Proportion of colon cancer attributable to lifestyle in a cohort of US women. Cancer Causes Control 2015;26:1271-9.
21. Nan H, Hutter CM, Lin Y et al. Association of aspirin and NSAID use with risk of colorectal cancer according to genetic variants. JAMA 2015;313:1133-42.
22. Liu Y, Tang W, Wang J et al. Association between statin use and colorectal cancer risk: a meta-analysis of 42 studies. Cancer Causes Control 2014;25:237-49.

23. Limsui D, Vierkant RA, Tillmans LS et al. Postmenopausal hormone therapy and colorectal cancer risk by molecularly defined subtypes among older women. Gut 2012; 61:1299-1305.
24. Ogino S, Chan AT, Fuchs CS et al. Molecular pathological epidemiology of colorectal neoplasia: an emerging transdisciplinary and interdisciplinary field. Gut 2011;60:397-411.
25. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell 2011;144: 646-74.
26. Colussi D, Brandi G, Bazzoli F et al. Molecular pathways involved in colorectal cancer: implications for disease behavior and prevention. Int J Mol Sci 2013;14:16365-85.
27. Grady WM, Carethers JM. Genomic and epigenetic instability in colorectal cancer pathogenesis. Gastroenterology 2008;135:1079-99.
28. Fearon, E. R. & Vogelstein, B. A genetic model for colorectal tumorigenesis. Cell 1990; 61: 759-767.
29. Lengauer C, Kinzler KW, Vogelstein B. Genetic instabilities in human cancers. Nature 1998;396:643-9.
30. Kinzler KW, Vogelstein B. Lessons from hereditary colorectal cancer. Cell 1996;87: 159-70.
31. Zeki SS, Graham TA, Wright NA. Stem cells and their implications for colorectal cancer. Nat Rev Gastroenterol Hepatol 2011;8:90-100.
32. Jones S, Chen WD, Parmigiani G et al. Comparative lesion sequencing provides insights into tumor evolution. Proc Natl Acad Sci USA 2008;105:4283-8.
33. Luo Y, Wong CJ, Kaz AM et al. Differences in DNA methylation signatures reveal multiple pathways of progression from adenoma to colorectal cancer. Gastroenterology 2014; 47:418-29. e8.
34. Van Engeland M, Derks S, Smits KM et al. Colorectal cancer epigenetics: complex simplicity. J Clin Oncol 2011;29:1382-91.
35. Bettington M, Walker N, Clouston A et al. The serrated pathway to colorectal carcinoma: current concepts and challenges. Histopathology 2013;62:367-86.
36. Rex DK, Ahnen DJ, Baron JA et al. Serrated lesions of the colorectum: review and recommendations from an expert panel. Am J Gastroenterol 2012;107:1315-29.
37. Kambara T, Simms LA, Whitehall VL et al. BRAF mutation is associated with DNA methylation in serrated polyps and cancers of the colorectum. Gut 2004;53:1137-44.
38. Noffsinger AE. Serrated polyps and colorectal cancer: new pathway to malignancy. Annu Rev Pathol 2009;4:343-64.
39. Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. Nature 2012;487:330-7.
40. Chittenden TW, Howe EA, Culhane AC et al. Functional classification analysis of somatically mutated genes in human breast and colorectal cancers. Genomics 2008;91: 508-11.
41. Parsons DW, Wang TL, Samuels Y et al. Colorectal cancer: mutations in a signalling pathway. Nature 2005 436:792.
42. Colussi D, Brandi G, Bazzoli F et al. Molecular pathways involved in colorectal cancer: implications for disease behavior and prevention. Int J Mol Sci 2013;14:16365-85.
43. Grady WM, Carethers JM. Genomic and epigenetic instability in colorectal cancer pathogenesis. Gastroenterology 2008;135:1079-99.
44. Bardelli A, Parsons DW, Silliman N et al. Mutational analysis of the tyrosine kinome in colorectal cancers. Science 2003;300:949.
45. Lao VV, Grady WM. Epigenetics and colorectal cancer. Nat Rev Gastroenterol Hepatol 2011;8: 686-700.
46. Kim YH, Petko Z, Dzieciatkowski S et al. CpG island methylation of genes accumulates during the adenoma progression step of the multistep pathogenesis of colorectal cancer. Genes Chromosomes Cancer 2006;45:781-9.
47. Bird A. The essentials of DNA methylation. Cell 1992;70:5-8.
48. Worthley D L, Whitehall VL, Buttenshaw RL et al. DNA methylation within the normal colorectal mucosa is associated with pathway-specific predisposition to cancer. Oncogene 2010; 29:1653-62.

49. Yamauchi M, Morikawa T, Kuchiba A et al. Assessment of colorectal cancer molecular features along bowel subsites challenges the conception of distinct dichotomy of proximal versus distal colorectum. Gut 2012;61:847-54.
50. Rosty C, Young JP, Walsh MD et al. PIK3CA activating mutation in colorectal carcinoma: associations with molecular features and survival. PLoS One 2013; 8: e65479.
51. Galon J, Mlecnik B, Bindea G et al. Towards the introduction of the "Immunoscore" in the classification of malignant tumours. J Pathol 2014;232:199-209.
52. Kostic AD, Gevers F, Pedamallu CS et al. Genomic analysis identifies association of Fusobacterium with colorectal carcinoma. Genome Res 2012;22:292-8.
53. Tahara T, Yamamoto E, Suzuki H et al. Fusobacterium in colonic flora and molecular features of colorectal carcinoma. Cancer Res 2014;74:1311-8.
54. Mima K, Sukawa Y, Nishihara R et al. Fusobacterium nucleatum and T cells in colorectal carcinoma. JAMA Oncol 2015;1:653-61.
55. Kuipers EJ, Grady WM, Lieberman D et al. Colorectal cancer . Nat Rev Dis Primers 2015;1:15065.
56. Moghimi-Dehkordi B, Safaee A, Zali MR. Prognostic factors in 1.138 Iranian colorectal cancer patients. Int J Colorectal Di 2008;23:683-8.
57. Yuan Y, Li MD, Hu HG et al. Prognostic and survival analysis of 837 Chinese colorectal cancer patients. World J Gastroenterol 2013;19:2650-9
58. Gan S, Wilson K, Hollington P. Surveillance of patients following surgery with curative intent for colorectal cancer. World J Gastroenterol 2007;13:3816-23.
59. Labianca R, Nordlinger B, Beretta GD et al. Early colon cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol 2013;Suppl 6:vi64-72.
60. Andrè T, Boni C, Navarro M et al. Improved overall survival with oxaliplatin, fluorouracil and leucovorin as adjuvant treatment in stage II or III colon cancer in the MOSAIC trial. J Clin Oncol 2009;27:3109-16.
61. Kuebler JP, Wieand HS, O'Connell MJ et al. Oxaliplatin combined with weekly bolus fluorouracil and leucovorin as surgical adjuvant chemotherapy for stage II and III colon cancer: results from NSABP C-07. J Clin Oncol 2007; 25:2198-2204.
62. Haller DG, Tabernero J, Maroun J et, al. Capecitabine plus oxaliplatin compared with fluorouracil and folinic acid as adjuvant therapy for stage III colon cancer. J Clin Oncol 2011;29:1465-71.
63. Moertel CG, O'Fallon JR, Go VL, et al. The preoperative carcinoembryonic antigen test in the diagnosis, staging, and prognosis of colorectal cancer. Cancer 1986;58:603-10.
64. Chen CS, Hsieh MC, Lai CC et al. Lead time of carcinoembryonic antigen elevation in the postoperative follo up of colorectal cancer did not affect the survival rate after recurrence. Int J Colorectal Dis 2010;25:567-71.
65. Tan E, Gouvas N, Nicholls RJ, et al. Diagnostic precision of carcinoembryonic antigen in the detection of recurrence of colorectal cancer. Surg Oncol 2009;1:15-24.
66. Diehl F, Schmidt K, Choti MA, Romans K, et al. Circulating mutanDNA to assess tumor dynamics. Nat Med 2008;14:985-90.
67. Inamura K. Colorectal cancers: An update on their molecular pathology. Cancers (Basel) 2018;10:26.
68. Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. Nature 2012;487:330-7.
69. Hao HX, Xie Y, Zhang Y et al. ZNRF3 promotes Wnt receptor turnover in an R-spondin-sensitive manner. Nature 2012;485,195-200.
70. Li Y, Oliver PG, Lu W et al. SRI36160 is a specific inhibitor of Wnt/beta-catenin signaling in human pancreatic and colorectal cancer cells. Cancer Lett. 2017; 389: 41-48.
71. Zhang B, Wang J, Wang X et al. Proteogenomic characterization of human colon and rectal cancer. Nature 2014;513:382-7.
72. Sadanandam A, Lyssiotis CA, Homicsko K et al. A colorectal cancer classification system that associates cellular phenotype and responses to therapy. Nat Med 2013;19: 619-25.

73. Chellappa K, Robertson GR, Sladek FM. HNF4alpha: A new biomarker in colon cancer? Biomark Med 2012;6:297-300.
74. Shimokawa T, Matsushima S, Tsunoda T et al. Identification of TOMM34, which shows elevated expression in the majority of human colon cancers, as a novel drug target. Int J Oncol 2006;29:381-6.
75. Irby RB, Mao W, Coppola D et al. Activating SRC mutation in a subset of advanced human colon cancers. Nat Genet 1999;21:187-90.
76. Budinska E, Popovici V, Tejpar S et al. Gene expression patterns unveil a new level of molecular heterogeneity in colorectal cancer. J Pathol 2013;231:63-76.
77. Marisa L, de Reynies A, Duval A et al. Gene expression classification of colon cancer into molecular subtypes: Characterization, validation, and prognostic value. PLoS Med 2013;10: e1001453.
78. Schlicker A, Beran G, Chresta C.M et al. Subtypes of primary colorectal tumors correlate with response to targeted treatment in colorectal cell lines. BMC Med Genom 2012;5:66.
79. Guinney J, Dienstmann R, Wang X et al. The consensus molecular subtypes of colorectal cancer. Nat Med 2015; 21:1350-6.
80. Sinicrope FA, Shi Q, Smyrk TC et al. Molecular markers identify subtypes of stage III colon cancer associated with patient outcomes. Gastroenterology 2015;148:88-99.
81. Sinicrope FA, Okamoto K, Kasi PM et al. Molecular biomarkers in the personalized treatment of colorectal cancer. Clin Gastroenterol Hepatol 2016;14:651-8.
82. Phipps AI, Limburg PJ, Baron JA et al. Association between molecular subtypes of colorectal cancer and patient survival. Gastroenterology 2015;148:77-87.
83. Pino MS, Chung DC. The chromosomal instability pathway in colon cancer. Gastroenterology 2010;138:2059-72.
84. Vogelstein B, Fearon ER, Hamilton SR et al. Genetic alterations during colorectal-tumor development. N Engl J Med 1988;319:525-32.
85. Lengauer C, Kinzler KW, Vogelstein B. Genetic instabilities in human cancers. Nature 1998;396:643-9.
86. Fodde R, Kuipers J, Rosenberg C et al. Mutations in the APC tumour suppressor gene cause chromosomal instability. Nat Cell Biol 2001;3:433-8.
87. Boland CR, Thibodeau SN, Hamilton SR et al. A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: Development of international criteria for the determination of microsatellite instability in colorectal cancer. Cancer Res 1998;58:5248-57.
88. Ogino S, Nosho K, Kirkner GJ et al. CpG island methylator phenotype, microsatellite instability, BRAF mutation and clinical outcome in colon cancer. Gut 2009;58:90-6.
89. Giannakis M, Hodis E, Jasmine Mu X et al. RNF43 is frequently mutated in colorectal and endometrial cancers. Nat Genet 2014;46:1264-6.
90. Yan HHN, Lai JCW, Ho SL et al. RNF43 germline and somatic mutation in serrated neoplasia pathway and its association with BRAF mutation. Gut 2017;66:1645-56.
91. Vilar E, Gruber SB. Microsatellite instability in colorectal cancer-the stable evidence. Nat Rev Clin Oncol 2010; 7:153-62.
92. Sargent DJ, Marsoni S, Monges G et al. Defective mismatch repair as a predictive marker for lack of efficacy of fluorouracil-based adjuvant therapy in colon cancer. J Clin Oncol 2010; 28:3219-26.
93. Collura A, Lagrange A, Svrcek M et al. Patients with colorectal tumors with microsatellite instability and large deletions in HSP110 T17 have improved response to 5-fluorouracil-based chemotherapy. Gastroenterology 2014; 146:401-11.
94. American Association for Cancer Research. First tissue-agnostic drug approval issued. Cancer Discov 2017; 7:656.
95. Gargiulo P, Della Pepa C, Berardi S et al. Tumor genotype and immune microenvironment in POLE-ultramutated and MSI-hypermutated Endometrial Cancers: New candidates for checkpoint blockade immunotherapy? Cancer Treat Rev 2016; 48:61-8.

96. Llosa NJ, Cruise M, Tam A et al. The vigorous immune microenvironment of microsatellite instable colon cancer is balanced by multiple counter-inhibitory checkpoints. Cancer Discov 2015; 5:43-51.
97. Jia M, Jansen L, Walter V et al. No association of CpG island methylator phenotype and colorectal cancer survival: Population-based study. Br J Cancer 2016; 115:1359-66.
98. Jia M, Gao X, Zhang Y et al. Different definitions of CpG island methylator phenotype and outcomes of colorectal cancer: A systematic review. Clin Epigenet 2016; 8:25.
99. East JE, Atkin WS, Bateman AC et al. British Society of Gastroenterology position statement on serrated polyps in the colon and rectum. Gut 2017; 66:1181-96.
100. Nishihara R, Wu K, Lochhead P et al. Long-term colorectal-cancer incidence and mortality after lower endoscopy. N. Engl J Med 2013; 369:1095-105.
101. Rodriguez-Soler M, Perez-Carbonell L, Guarinos C et al. Risk of cancer in cases of suspected lynch syndrome without germline mutation. Gastroenterology 2013; 144:926-32.
102. Briggs S, Tomlinson I. Germline and somatic polymerase epsilon and delta mutations define a new class of hypermutated colorectal and endometrial cancers. J Pathol 2013; 230:148-53.
103. Palles C, Cazier JB, Howarth KM et al. D Germline mutations affecting the proofreading domains of POLE and POLD1 predispose to colorectal adenomas and carcinomas. Nat Genet 2013; 45:136-44.
104. Stenzinger A, Pfarr N, Endris V et al. Mutations in POLE and survival of colorectal cancer patients-Link to disease stage and treatment. Cancer Med 2014; 3:1527-38.
105. Dudley JC, Lin MT, Le DT et al. Microsatellite Instability as a Biomarker for PD-1 Blockade. Clin Cancer Res 2016; 22:813-20.
106. Gong J, Wang C, Lee PP et al. Response to PD-1 Blockade in Microsatellite Stable Metastatic Colorectal Cancer Harboring a POLE Mutation. J Natl Compr Cancer Netw 2017; 15:142-7.
107. Douillard JY, Oliner KS, Siena S et al. Panitumumab-FOLFOX4 treatment and RAS mutations in colorectal cancer. N Engl J Med 2013; 369:1023-34.
108. Takane K, Akagi K, Fukuyo M et al. DNA methylation epigenotype and clinical features of NRAS-mutation (+) colorectal cancer. Cancer Med 2017; 6:1023-35.
109. De Roock W, Claes B, Bernasconi D et al. 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 2010; 11:753-62.
110. Van Brummelen EMJ, de Boer A, Beijnen JH et al. BRAF mutations as predictive biomarker for response to anti-EGFR monoclonal antibodies. Oncologist 2017;22.864-72.
111. Xu JM, Wang Y, Wang YL et al. PIK3CA Mutations Contribute to Acquired Cetuximab Resistance in Patients with Metastatic Colorectal Cancer. Clin Cancer Res 2017; 23:4602-16.
112. Jones RP, Sutton PA, Evans JP et al. Specific mutations in KRAS codon 12 are associated with worse overall survival in patients with advanced and recurrent colorectal cancer. Br J Cancer 2017; 116:923-9.
113. Sinicrope FA, Mahoney MR, Yoon HH et al. Analysis of Molecular Markers by Anatomic Tumor Site in Stage III Colon Carcinomas from Adjuvant Chemotherapy Trial NCCTG N0147 (Alliance). Clin Cancer Res 2015; 21:5294-304.
114. Taieb J, Le Malicot K, Shi Q et al. Prognostic Value of BRAF and KRAS Mutations in MSI and MSS Stage III Colon Cancer. J Natl Cancer Inst 2016;109 (5).
115. Mouradov D, Domingo E, Gibbs P et al. Survival in stage II/III colorectal cancer is independently predicted by chromosomal and microsatellite instability, but not by specific driver mutations. Am J Gastroenterol 2013; 108:1785-93.
116. Summers MG, Smith CG, Maughan TS et al. BRAF and NRAS locus-specific variants have different outcomes on survival to colorectal cancer. Clin Cancer Res 2017; 23:2742-9.
117. Corcoran RB, Ebi H, Turke AB et al. EGFR-mediated re-activation of MAPK signaling contributes to insensitivity of BRAF mutant colorectal cancers to RAF inhibition with vemurafenib. Cancer Discov 2012; 2:227-35.

118. Van Cutsem E, Huijberts S, Grothey A et al. Binimetinib, Encorafenib, and Cetuximab triplet therapy for patients with BRAF V600E-Mutant metastatic colorectal cancer: safety lead-in results from the Phase III BEACON Colorectal Cancer Study. J Clin Oncol 2019; doi: 10.1200/JCO.18.02459.
119. Lee DW, Han SW, Cha Y et al. Association between mutations of critical pathway genes and survival outcomes according to the tumor location in colorectal cancer. Cancer 2017; 123:3513-23.
120. Chan AT, Ogino S, Fuchs CS. Aspirin and the risk of colorectal cancer in relation to the expression of COX-2. N Engl J Med 2007; 356:2131-42.
121. Liao X, Lochhead P, Nishihara R et al. Aspirin use, tumor PIK3CA mutation, and colorectal-cancer survival. N Engl J Med 2012; 367:1596-606.
122. Gray RT, Cantwell MM, Coleman HG et al. Evaluation of PTGS2 expression, PIK3CA mutation, aspirin use and colon cancer survival in a population-based cohort study. Clin Transl Gastroenterol 2017;8:e91.
123. Hua X, Phipps AI, Burnett-Hartman AN et al. Timing of aspirin and other nonsteroidal anti-inflammatory drug use among patients with colorectal cancer in relation to tumor markers and survival. J Clin Oncol 2017; 35:2806-13.
124. Nan H, Hutter CM, Lin Y et al. Association of aspirin and NSAID use with risk of colorectal cancer according to genetic variants. JAMA 2015; 313:1133-42.
125. Fink SP, Yamauchi M, Nishihara R et al. Aspirin and the risk of colorectal cancer in relation to the expression of 15-hydroxyprostaglandin dehydrogenase (HPGD). Sci Transl Med 2014; 6:233re2.
126. Lengauer C, Kinzler KW, Vogelstein B. Genetic instabilities in human cancers. Nature 1998; 396:643-9.
127. Derynck R, Akhurst RJ, Balmain A. TGF-beta signaling in tumor suppression and cancer progression. Nat Genet 2001; 29:117-29.
128. Liu J, Zhang C, Hu W et al. Tumor suppressor p53 and its mutants in cancer metabolism. Cancer Lett 2015; 356:197-203.
129. Le DT, Durham JN, Smith KN et al. Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. Science 2017; 357:409-13.
130. Overman MJ, McDermott R, Leach JL et al. Nivolumab in patients with metastatic DNA mismatch repair-deficient or microsatellite instability-high colorectal cancer (CheckMate 142): An open-label, multicentre, phase 2 study. Lancet Oncol 2017;18: 1182-91.
131. Gentles AJ, Newman AM, Liu CL et al. The prognostic landscape of genes and infiltrating immune cells across human cancers. Nat Med 2015; 21:938-45.
132. Giannakis M, Mu XJ, Shukla SA et al. Genomic Correlates of Immune-Cell Infiltrates in Colorectal Carcinoma. Cell Rep 2016; 15:857-65.
133. Masugi Y, Nishihara R, Yang J et al. Tumour CD274 (PD-L1) expression and T cells in colorectal cancer. Gut 2017; 66:1463-73.
134. Dalerba P, Sahoo D, Paik S et al. CDX2 as a prognostic biomarker in stage II and stage III colon cancer. N Engl J Med 2016; 374:211-22.
135. Bertotti A, Migliardi G, Galimi F et al. A Molecularly annotated platform of patient-derived xenografts (‘Xenopatients’) identifies HER2 as an effective therapeutic target in cetuximab-resistant colorectal cancer. Cancer Discov 2011; 1:508-23.
136. Sartore-Bianchi A, Trusolino L, Martino C et al. Dual-targeted therapy with trastuzumab and lapatinib in treatment-refractory, KRAS codon 12/13 wild-type, HER2-positive metastatic colorectal cancer (HERACLES): a proof-of-concept, multicentre, open-label, phase 2 trial. Lancet Oncol 2016; 17:738-46.
137. Murtaza M, Dawson SJ, Tsui DW, et al: Non-invasive analysis of acquired resistance to cancer therapy by sequencing of plasma DNA. Nature 2013; 497:108-12.
138. Diaz LA Jr, Williams RT, Wu J, et al: The molecular evolution of acquired resistance to targeted EGFR blockade in colorectal cancers. Nature 2012; 486:537-40.

139. Diaz LA Jr, Bardelli A. Liquid biopsies: genotyping circulating tumor DNA. J Clin Oncol 2014; 32:579-86.
140. Mandel P, Metais P: Les acides nucleiques du plasma sanguin chez l'homme. C R Seances Soc Biol Fil 1948; 142:241-3.
141. Leon SA, Shapiro B, Sklaroff DM et, al: Free DNA in the serum of cancer patients and the effect of therapy. Cancer Res 1977, 37:646-50.
142. Siravegna G, Bardelli A. Genotyping cell-free tumor DNA in the blood to detect residual disease and drug resistance.Genome Biol 2014; 15(8):449.
143. Diehl F, Schmidt K, Choti MA, Romans K, et al. Circulating mutanDNA to assess tumor dynamics. Nat Med 2008; 14:985-90
144. Bettegowda C, Sausen M, Leary RJ, et al. Detection of circulating tumor DNA in early- and late-stage human malignancies. Sci TranslMed 2014; 6:224ra24.
145. Garcia-Murillas I, Schiavon G, Weigelt B, et al. Mutation tracking in circulating tumor DNA predicts relapse in early breast cancer. Sci Transl Med 2015; 7:302ra133.
146. Tie J, Wang Y, Tomasetti C, et al. Circulating tumor DNA analysis detects minimal residual disease and predicts recurrence in patients with stage II colon cancer. Sci Transl Med 2016; 8:346ra92.
147. Diaz LA Jr, Williams RT, Wu J, et al: The molecular evolution of acquired resistance to targeted EGFR blockade in colorectal cancers. Nature 2012; 486:537-40.
148. Forshew T, Murtaza M, Parkinson C et al. Noninvasive identification and monitoring of cancer mutations by targeted deep sequencing of plasma DNA. Sci Transl Med 2012; 4:136ra68.
149. Calon A, Espinet E, Palomo-Ponce S, et al. Dependency of colorectal cancer on a TGF-β-driven program in stromal cells for metastasis initiation. Cancer Cell 2012; 22:571-84.
150. Knüpfer H, Preiss R. Serum interleukin-6 levels in colorectal cancer patients--a summary of published results. Int J Colorectal Dis 2010; 25:135-40.
151. Putoczki TL, Thiem S, Loving A, et al. Interleukin-11 is the dominant IL-6 family cytokine during gastrointestinal tumorigenesis and can be targeted therapeutically. Cancer cell 2013; 24:257-71.
152. Cohen JD, Li L, Wang Y, et al. Detection and localization of surgically resectable cancers with a multi-analyte blood test. Science 2018; 359:926-30.
153. Wang Y, Li L, Cohen JD, et al. Prognostic potential of circulating tumor DNA measurement in postoperative surveillance of nonmetastatic colorectal cancer. JAMA Oncol 2019. doi: 10.1001/jamaoncol.2019.0512.
154. Reinert T, Henriksen TV, Christensen E, et al. Analysis of plasma cell-free DNA by ultradeep sequencing in patients with stages I to III colorectal cancer. JAMA Oncol 2019. doi: 10.1001/jamaoncol.2019.0528.
155. Calon A, Espinet E, Palomo-Ponce S, et al. Dependency of colorectal cancer on a TGF-β-driven program in stromal cells for metastasis initiation. Cancer Cell. 2012;22(5):571-84.
156. Van den Eynde M, Mlecnik B, Bindea G The Link between the Multiverse of Immune Microenvironments in Metastases and the Survival of Colorectal Cancer Patients. Cancer cell 2018;34:1012-26.
157. Galon J, Mlecnik B, Bindea G, et al. Towards the introduction of the 'Immunoscore' in the classification of malignant tumours. J Pathol 2014;232,199-209.
158. Elliott RL, Blobe GC. Role of transforming growth factor beta in human cancer. J Clin. Oncol 2005;23:2078-93.
159. Massague J, Blain SW, Lo RS. TGFb signaling in growth control, cancer, and heritable disorders. Cell 2000; 103:295-309.
160. Robson H, Anderson E, James RD, et al. Transforming growth factor beta 1 expression in human colorectal tumours: an independent prognostic marker in a subgroup of poor prognosis patients. Br J Cancer 1996;74:753-8.
161. Tsushima H, Ito N, Tamura S, et al. Circulating transforming growth factor beta 1 as a predictor of liver metastasis after resection in colorectal cancer. Clin Cancer Res 2001;7:1258-62.

162. Massague J. TGFbeta in Cancer. Cell 2008;134: 215-30.
163. Tauriello DF, Batlle E. Targeting the microenvironment in advanced colorectal cancer. Trends Cancer 2016; 2:495-504
164. Fujii M, Shimokawa M, Date S, et al. A colorectal tumor organoid library demon-strates progressive loss of niche factor requirements during tumorigenesis. Cell Stem Cell 2016;18: 827-38.
165. Yan P, Klingbiel D, Saridaki Z, et al. Reduced expression of SMAD4 is associated with poor survival in colon cancer. Clin Cancer Res 2016;22:3037-47.
166. Fessler E, Drost J, van Hooff SR, et al. TGFβ signaling directs serrated adenomas to the mesenchymal colorectal cancer subtype. EMBO Mol Med 2016;8:745-60.
167. Markowitz SD, Bertagnolli MM. Molecular origins of cancer: molecular basis of colorectal cancer. N Engl J Med 2009;361:2449-60.
168. Muñoz NM, Upton M, Rojas A, et al. Transforming growth factor beta receptor type II inactivation induces the malignant transformation of intestinal neoplasms initiated by Apc mutation. Cancer Res 2006;66:9837-44.
169. Wang J, Sun L, Myeroff L, et al. Demonstration that mutation of the type II transforming growth factor beta receptor inactivates its tumor suppressor activity in replication error-positive colon carcinoma cells. J Biol Chem 1995; 270:22044-9.
170. Becker C, Fantini MC, Schramm C, et al. TGF-β suppresses tumor progression in colon cancer by inhibition of IL-6 trans-signaling. Immunity 2004;21:491-501.
171. Kopf M, Baumann H, Free G, et al. Impaired immune and acute-phase responses in interleukin-6-deficient mice. Nature 1994;368:339-42.
172. Kumari N, Dwarakanath BS, Das A, et al. Role of interleukin-6 in cancer progression and therapeutic resistance. Tumour Biol 2016;37:11553-72.
173. Calon A, Lonardo E, Berenguer-Llergo A, et al. Stromal gene expression defines poor-prognosis subtypes in colorectal cancer. Nat Genet 2015;47:320-9.
174. de Gramont A, Faivre S, Raymond E. Novel TGF-β inhibitors ready for prime time in onco-immunology. Oncoimmunology 2016;6:e1257453.
175. Jones SA, Horiuchi S, Topley N, et al. The soluble interleukin 6 receptor: mechanisms of production and implications in disease. FASEB J 2001;15:43-58.
176. Scheller J, Chalaris A, Schmidt-Arras D, et al. The pro- and anti-inflammatory properties of the cytokine interleukin-6. Biochim Biophys Acta 2011;1813:878-88.
177. Bournazou E, Bromberg J. Targeting the tumor microenvironment: JAK-STAT3 signaling. JAKSTAT 2013;2:e23828.
178. Campbell IL, Erta M, Lim SL, et al. Trans-signaling is a dominant mechanism for the pathogenic actions of interleukin-6 in the brain. J Neurosci 2014;34:2503-13.
179. Wolf J, Rose-John S, Garbers C. Interleukin-6 and its receptors: a highly regulated and dynamic system. Cytokine 2014;70:11-20.
180. Zhang HF, Chen Y, Wu C, et al. The opposing function of STAT3 as an oncoprotein and tumor suppressor is dictated by the expression status of STAT3beta in esophageal squamous cell carcinoma. Clin Cancer Res 2016;22:691-703.
181. Subramaniam KS, Omar IS, Kwong SC, et al. Cancer-associated fibroblasts promote endometrial cancer growth via activation of interleukin-6/STAT-3/c-Myc pathway. Am J Cancer Res 2018;6:200-13.
182. Spitzner M, Ebner R, Wolff HA, et al. STAT3: a novel molecular mediator of resistance to chemoradiotherapy. Cancers (Basel) 2014;6:1986-2011.
183. Chen R, Chen B. Siltuximab (CNTO 328): a promising option for human malignancies. Drug des Devel Ther 2015; 9:3455-8.
184. Ham IH, Oh HJ, Jin H, et al. Targeting interleukin-6 as a strategy to overcome stroma-induced resistance to chemotherapy in gastric cancer. Mol cancer 2019; 18:68.
185. Gordon MS, McCaskill-Stevens WJ, Battiato LA, et al. A phase I trial of recombinant human interleukin-11 (neumega rhIL-11 growth factor) in women with breast cancer receiving chemotherapy. Blood 1996;87:3615-24.

186. Putoczki TL, Thiem S, Loving A, et al. R.A. Interleukin-11 is the dominant IL-6 family cytokine during gastrointestinal tumorigenesis and can be targeted therapeutically. Cancer Cell 2013;24:257-71.
187. Matadeen R, Hon WC, Heath JK, et al. The dynamics of signal triggering in a gp130-receptor complex. Structure 2007;15:441-8.
188. Babon JJ, Varghese LN, Nicola NA. Inhibition of IL-6 family cytokines by SOCS3, Semin Immunol 2014;26:13-9.
189. Thiem S, Pierce TP, Palmieri M, et al. mTORC1 inhibition restricts inflammation-associated gastrointestinal tumorigenesis in mice. J Clin Invest 2013;123:767-81.
190. Kaye JA. Clinical development of recombinant human interleukin-11 to treat chemotherapy-induced thrombocytopenia, Curr Opin Hematol 1996;3:209-15.
191. Cohen JD, Javed AA, Thoburn C, et al. Combined circulating tumor DNA and protein biomarker-based liquid biopsy for the earlier detection of pancreatic cancers. Proc Natl Acad Sci USA 2017;114(38):10202-7.
192. Smith T, Heger A, Sudbery I. UMItools: modeling sequencing errors in Unique Molecular Identifiers to improve quantificationaccuracy. Genome Res 2017;27:491-9.
193. Schimieder R, Edwards R. Quality control and preprocessing of metagenomic datasets. Bioinformatics 2011; 27:863-4.
194. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 2009;25:1754-60.
195. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics 2009;25:2078-9.
196. Au CH, Ho DN, Kwong A, et al. BAMClipper: removing primers from alignments to minimize false-negative mutations in ampliconnext-generation sequencing. Sci Rep 2017;7:1567.
197. Cibulskis K, Lawrence MS, Carter SL, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nat Biotechnol 2013;31:213-9.
198. Lai Z, Markovets A, Ahdesmaki M, et al. VarDict: a novel and versatile variant caller for next-generation sequencing in cancer research. Nucleic Acids Res 2016;44:e108.
199. Xu C, Nezami Ranjbar MR, Wu Z, et al. Detecting very low allele fraction variants using targeted DNA sequencing and a novel molecular barcode-aware variant caller. BMC Genomics 2017;18:5.
200. Danecek P, Auton A, Abecasis G, et al. The variant call format and VCFtools. Bioinformatics 2011;27:156-8.
201. Cingolani P, Patel VM, Coon M, et al. Using Drosophila melanogaster as a Model for Genotoxic Chemical Mutational Studies with a New Program, SnpSift. Front Genet 2012;3:35.
202. McLaren W, Gil L, Hut SE, et al. The Ensembl Variant Effect Predictor. Genome Biol 2016;17:122.
203. Erdmann J. All aboard: Will molecular tumor boards help canvcer patients?. Nat Med 2015;21:655-6.
204. Dienstmann R, Dong F, Borger D, et al. Standardized decision support in next generation sequencing reports of somatic cancer variants. Mol Oncol 2014;8:859-73.
205. Lee LA, Arvai KJ, Jones D. Annotation of sequence variants in cancer samples: processes and pitfalls for routine assays in the clinical laboratory. J Mol Diagn 2015;17:339-51.
206. Froyen G, Broekmans A, Hillen F, et al. Validation and application of a custom-designed targeted Next-Generation Sequencing panel for the diagnostic mutational profiling of solid tumors. PloS One 2016;11:e0154038.
207. Vogelstein B, Kinzler KW. Digital PCR. Proc Natl Acad Sci U S A 1999;96:9236-41.
208. Diehl F, Schmidt K, Choti MA, et al. Circulating mutant DNA to assess tumor dynamics. Nat Med 2008;14:985-90.
209. Tie J,Cohen JD,Wang Y, et al. Serial circulating tumour DNA analysis during multimodality treatment of locally advanced rectal cancer: a prospective biomarker study. Gut 2019;68:663-71.
210. Coakley M, Garcia-Murillas I, Turner NC. Molecular residual disease and adjuvant trials design in solid tumors. Clin Cancer Res 2019; doi: 10.1158/1078-0432.CCR-19-0152.

211. Shim KS, Klim KH, Han Ws, et al. Elevated serum levels of transforming growth factor-beta1 in patients with colorectal carcinoma: its association with tumor progression and its significant decrease after curative surgical resection. Cancer 1999;85:554-61.
212. Ragulan C, Eason K, Fontana E, et al. Analytical validation of multiplex biomarker assay to stratify colorectal cancer into molecular subtypes. Sci Rep 2019; doi.org/10.1038/s41598-019-43492-0.
213. Van Cutsem E, Labianca R, Bodoky G, et al. Randomized phase III trial comparing biweekly infusional fluorouracil/leucovorin alone or with iri- notecan in the adjuvant treatment of stage III colon cancer: PETACC-3. JCO 2009; 27:3117-25.
214. ong N, Pogue-Geile KL, Gavin PG, et al. Clinical outcome from oxalipla-tin treatment in stage II/III colon cancer according to intrinsic subtypes: secondary analysis of NSABP C-0 7/NRG oncology randomized clinical trial. JAMA Oncol 2016;2:1162-9.
215. Alderdice M, Richman SD, Gollins S, et al. Prospective patient stratification into robust cancer cell intrinsic subtypes from colorectal cancer biopsies. J Pathol 20 18;245:19-28.
216. Fontana E, Eason K, Cervantes A, et al. Context matters-consensus molecular subtypes of colorectal cancer as biomarkers for clinical trials. Ann Oncol 2019;30:520-7.
217. Pilati C, Taieb J, Balogoun R, et al. CDX2 prognostic value in stage II/III resected colon cancer is related to CMS classification. Ann Oncol 2017;28:1032-5.
218. Mei ZB, Duan CY, Li CB, et al. Prognostic role of tumor PIK3CA mutation in colorectal cancer: a systematic review and meta-analysis. Ann oncol 2016; 27:1836-48.
219. Therkildsen C, Bergmann TK, Henrichsen-Schnack T, et al. The predictive value of KRAS, NRAS, BRAF, PIK3CA and PTEN for anti-EGFR treatment in metastatic colorectal cancer: a systematic review and meta-analysis. Acta Oncol 2014;53:852-64.

There are no more than two rules for writing: having something to say, and saying it
Oscar Wilde