UNIVERSITY OF NAPLES FEDERICO II

DOCTORATE IN

MOLECULAR MEDICINE AND MEDICAL BIOTECHNOLOGY

XXXIV CYCLE

New dimension of cell culture in physiopathology: development of patient-derived organoid structure and analysis in colorectal cancer

Tutor
Prof. Lucio Pastore

Candidate
Federica Di Maggio

Co-Tutor
Prof. Francesco Salvatore

2018-2022

Table of Contents

List of abbreviations

AD: Autosomal Dominant

Advanced Dmem/F12: Advanced Dulbecco's modified Eagle medium/ F12

AJCC: American Joint Committee on Cancer

ASCs: Adult Stem Cells

CIMP: CpG island methylator phenotype

CIN: Chromosomal instability

CNVs: Copy number variants

CRC: Colorectal Cancer

CTC: Computed tomography colonography

EDTA: Ethylene Diamine Tetra Acetic acid

FAP: Familial Adenomatous Polyposis

FIT: Fecal immunochemical test

FOBT: Fecal occult blood test

gDNA: genomic DNA

HNPCC: Hereditary Non-polyposis Colorectal Cancer

IBD: Inflammatory bowel disease

JPS: Juvenile polyposis syndrome

LS: Lynch syndrome

MAP: MUTYH-associated polyposis

MMR: Mismatch repair

MMR-d: DNA mismatch repair deficient

MMR-p: DNA mismatch repair proficient

MSI: DNA microsatellites instability

NGS: Next Generation Sequencing

ONT: Oxford nanopore technology

PBS: Phosphate buffered saline

PDOs: Patient-derived Organoids

PDTOs: Patient derived tumor organoids

PDTX: Patient-derived tumor xenograft

PF: Clusters passing filter

PSCs: Pluripotent stem cells

Q-score: Quality score

SNV: Single nucleotide variant

TNM: Tumor/node/metastasis classification

VCF: Variant call format

WGS: Whole Genome Sequencing

Figures Index

Abstract

Colorectal cancer (CRC) is the fourth most frequent malignancy and the third leading cause of cancer death worldwide. About 5-10% of patients affected by CRC carry pathogenic germline variants in genes associated with cancer risk development. This suggests that the mutational status in predisposing genes may be a tool to identify at risk subjects; these can be considered good candidates for targeted prevention and/or treatment. In this context, the establishment of patient-derived organoids (PDOs) represent a milestone toward a successful precision medicine. The aim of this study was to establish, and then to determine, whether PDOs reflect the genes molecular pattern of tumoral tissue of CRC patients and so identify suitable candidates for treatment. We enrolled eighty patients affected by CRC who underwent surgery and, from these we stabilized 11 PDOs. They were monitored during the various phases of growth using an advanced microscopy instrumentation (Cell Discoverer-7, Zeiss). Moreover, we carried out molecular analyses, first using customized multi-gene panel, and then using whole genome sequencing (WGS) with Oxford Nanopore Technology (ONT) to evaluate the gene pattern of four genomes derived from the same patient (blood, PDOs, tumor-derived tissue and paired healthy tissue). The established model system showed genomic sequencing concordance with paired tumor in the same patient. Indeed, we confirmed the presence of 37 pathogenic mutations at somatic level in the PDOs. With these experiments we indicate a robust validity of the strategy used, also to pursue with further experimentation toward the precision medicine goal. In conclusion, the full spectrum of mutations to the aim of precision medicine studies, as also demonstrated by this work, includes the possibility of subsequent screening of cancer drug libraries, then leading to targeting specific mutations.

1. Introduction

1.1 Cancer type and its development

In normal conditions, the growth of cells is usually organized by different control mechanisms within the cell. On the other hand, cancer can be defined as a complex human disease where a group of abnormal cells grow uncontrollably; these cells can infiltrate the normal organs and tissues of the body, altering their structure and function. The tumor cells, in emerging neoplastic clone, accumulate a series of genetic and epigenetic alterations that tend to modify activities of several genes and their products causing various phenotypic changes. Moreover, malignant tumors invade nearby tissue borders and have the capacity to infiltrate the adjacent structures up to distant organs creating metastases; contrary, benign tumors remain localized and are usually manageable by surgery.

Over the years, to better understand tumor biology, “cancer hallmarks” have been identified. Hanahan and Weinberg [1] revisited the original distinctive hallmarks, thus, identifying 10 in total instead of 6 (Figure 1). All these features contribute to the initiation and the progression of tumorigenesis, and it takes place in a multistep sequence. This process can be caused by internal factors such as heritable predisposing mutations, in 5-10% of the cases, replication errors, hormones and immune conditions. The presence of germline mutations can predispose to the hereditary and familial cancer such as breast, ovarian and colorectal cancer (CRC).

Due to the big complexity of this disease, the possibility of developing a completely precision/personalized medicine direct to specific existing mutations, also with the help of predictive medicine, and through the sequencing of the human

genome has become more and more a challenge and a goal in the treatment of tumors.

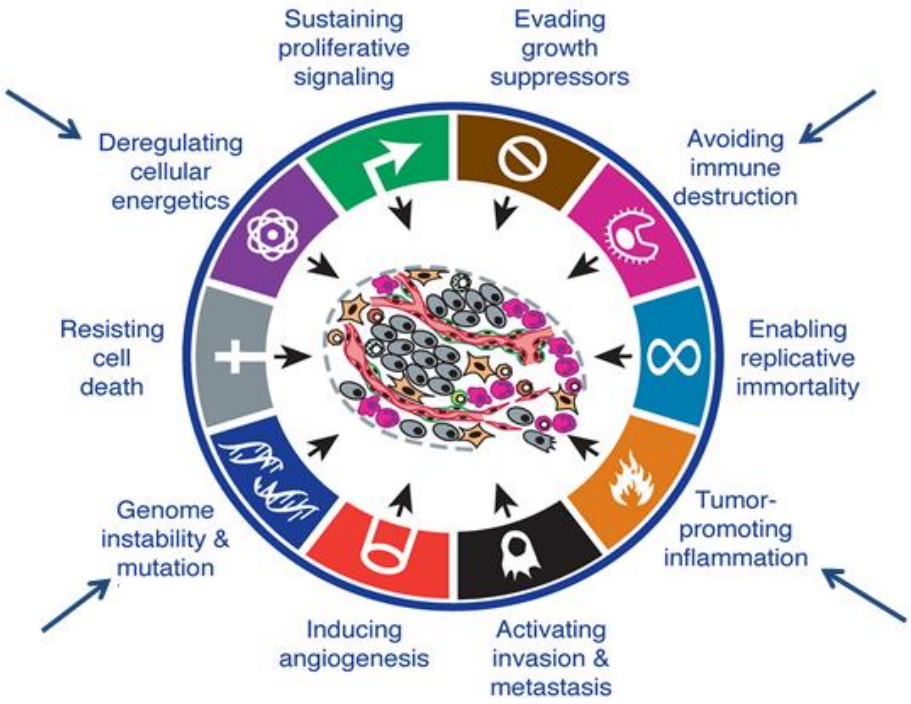


Figure 1. The ten hallmarks of cancer. All the hallmarks are briefly described and with the blue arrows are identified the 4 new ones. Figure modified by Hanahan and Weinberg [1].

Indeed, the predictive medicine can be useful, and the treatment can be effective based on the mutational pattern of each patient; to date, for breast, ovarian, colorectal, and gastric cancer, different treatments are available in the field of precision/personalized medicine.

1.2 Epidemiology of cancers, focus on Colorectal Cancer (CRC)

Based on the population data collected in 2020, the incidence of cancer worldwide accounts for 19 million (Figure 2), in Italy it accounts for 415.269; for the death of

cancer worldwide it accounts for nearly 10 million, instead, in Italy are estimated 174.759 deaths [2]. In Italy, the most frequently diagnosed cancer is breast (13.3% of all new diagnoses), followed by colorectal (11.7%), lung (10.1%), prostate (9.5%) and bladder (6.8%) cancer.

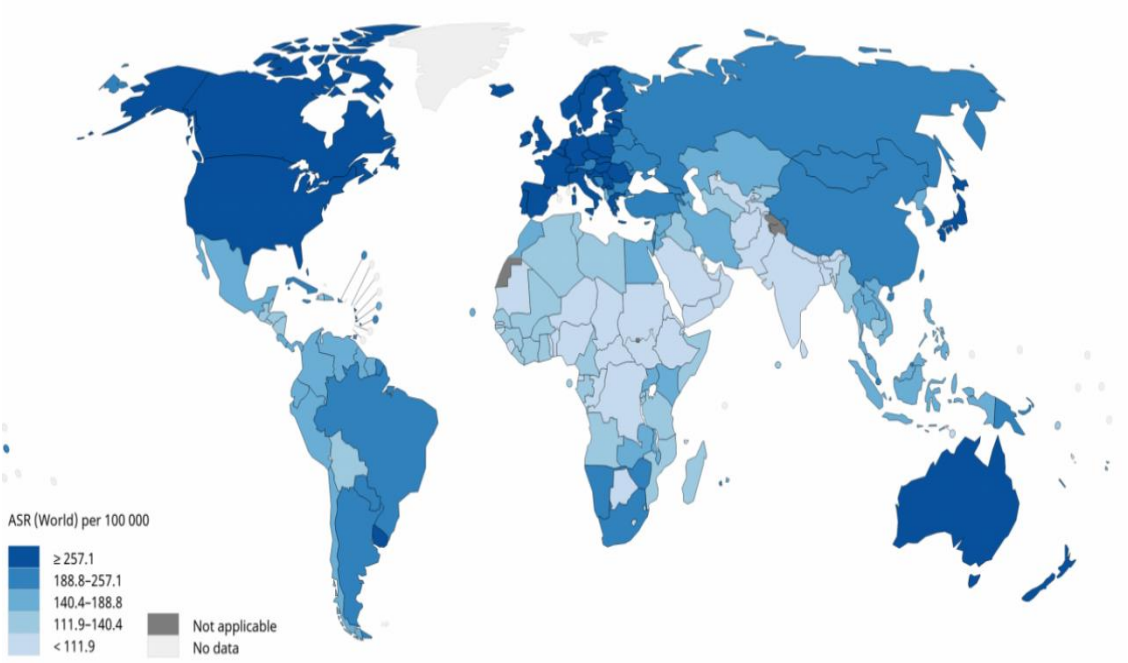
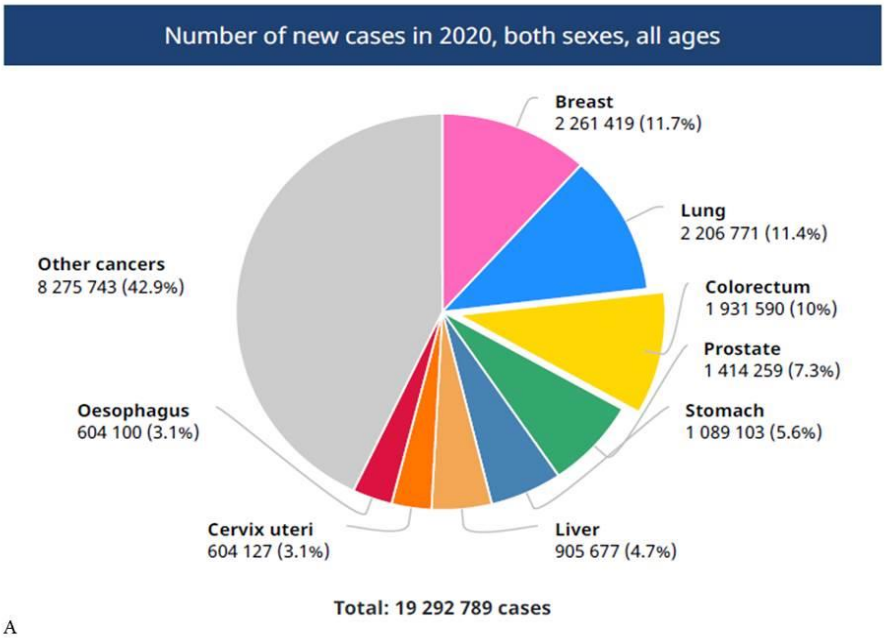
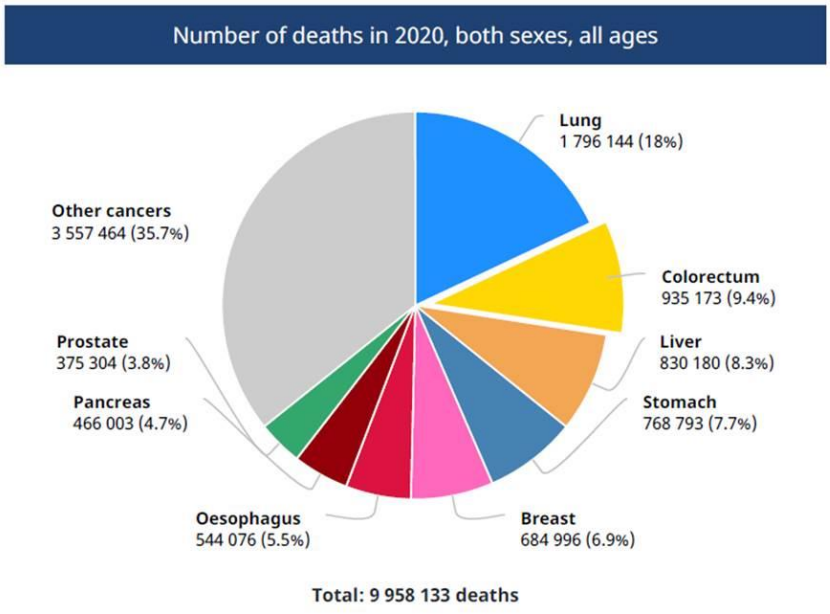


Figure 2. Estimated age-standardized incidence rates worldwide in 2020 of all cancers, both sexes, all ages. Data available from: http://globocan.iarc.fr [2].

Among the common malignancies, CRC is the second leading cause of cancer-related death in man and the third most common cancer worldwide in both sex; it accounts approximately for 1.9 million new cases and 935,173 deaths in 2020 worldwide [2-3]. By 2030, CRC is estimated to rise with more than 2.2 million new cases and 1.1 million cancer deaths [4-5]. CRC incidence and mortality rates vary widely across the world (Figure 3), in Italy the incidence is 11,6%, among all the new diagnoses, for both sexes.



A



B

Figure 3. Worldwide colorectal cancer epidemiology. A) Number of new cancer cases, in both sexes and in all ages; CRC is the third one. B) Number of deaths in 2020, in both sexes and in all ages; CRC is the second leading cause, and it accounts for 9.4% of cancer-related death. Data available from: http://globocan.iarc.fr [2].

In the last years, it has been an overall decrease in CRC incidence and mortality among individuals age 50 and older but recent epidemiological studies demonstrate increasing incidence of CRC among young individuals which remain unexplained [2].

Of all CRC cases, genetic predisposition, due to pathogenic germline variants in genes associated with cancer risk development, accounts for about 5-10% [6], of all CRC cases. In addition, the incidence is increasing rapidly in many low-income and middle-income countries, particularly those undergoing social and economic changes. In several high-income countries such as USA, CRC incidence and mortality have been diminishing or stabilizing because of improvements in early detection, prevention, perioperative care, as well as chemotherapy and radiotherapy techniques [7-8]. CRC survival rates decrease with age and are highly dependent mainly with the stage of the disease at diagnosis: 5-years survival rates range from 90% for stage I to 14% for advanced stage IV with distant metastases [9].

1.3 Colorectal anatomy

To make out colorectal cancer, it is helpful to understand what parts of the body are affected and how they work. The colon (large intestine) is the distal part of the gastrointestinal tract, extending from the cecum to the anal canal. It receives digested food from the small intestine, from which it absorbs water and electrolytes and form faeces [10-11].

The colon averages 150 cm in length and can be divided into four parts (proximal to distal): ascending, transverse, descending and sigmoid. Then, follow the sigmoid colon there is the rectum which then ends in the anal canal.

The ascending colon (right-sided) is where undigested food begins its journey through the colon. Undigested food moves upwards through this section, where fluid is reabsorbed more efficiently. When it meets the right lobe of the liver, it turns 90 degrees to move horizontally. This turn is known as the right colic flexure (or hepatic flexure) and marks the start of the transverse colon. Moving across the body, the **transverse colon** takes the food from one side of the body to the other (right to left). The transverse colon extends from the right colic flexure to the spleen, where it turns another 90 degrees to point down. This turn is known as the left colic flexure (or splenic flexure). Once the food has travelled across the top through the transverse colon, it makes its way downward through the **descending colon**, typically on the left side. After the left colic flexure, the colon moves inferiorly towards the pelvis. The final section of the colon is called **sigmoid colon,** this portion is shaped like an "S" and it is the last step before the rectum. **The rectum** is a 5 to 6-inch chamber that connects the colon to the anus. It is the job of the rectum to act as a reservoir unit and holds the stool until defecation (evacuation) occurs [12-13] (Figure 4A).

The colon has the typical histological structure as the digestive tube: *mucosa, submucosa, muscularis and serosa/adventitia* (Figure 4B). The mucosa is lined by simple columnar epithelium (*lamina epithelialis*) with long microvilli. It is covered by a layer of mucus which aids the transport of the feces [14]. The mucosa does not contain villi but many crypts of Lieberkuhn in which numerous goblet cells and enteroendocrine cells are found.

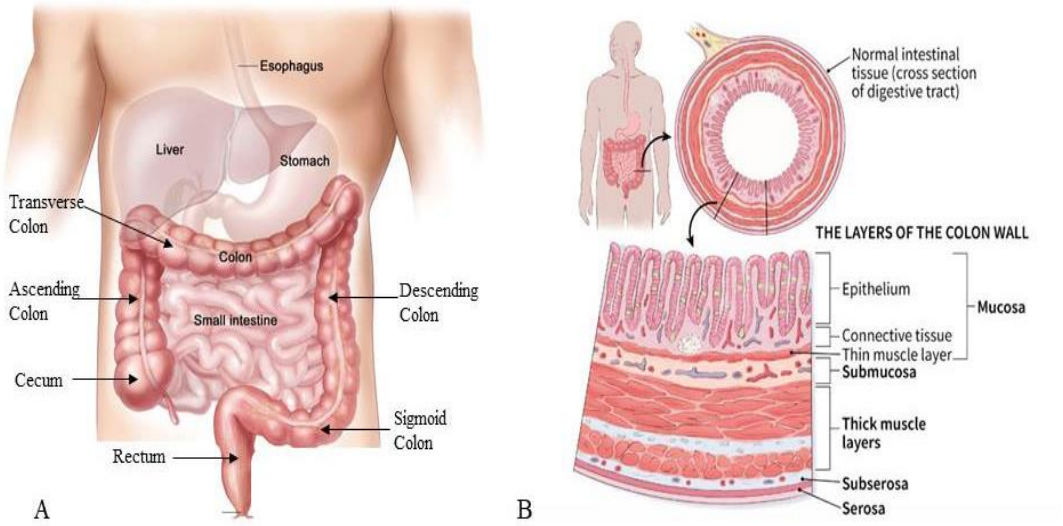


Figure 4. Anatomy and Histology of the Colon. A) All the different anatomical features of the colon are shown. Starting from the cecum to the anal canal. B) Starting from the inner part the layers of the colon are: mucosa, submucosa, muscularis and serosa/adventitia. Figure by American Cancer Society [15].

1.4 Risk factors and genetic causes

The etiology of CRC is complex and multi-causal and is associated with both modifiable and non-modifiable factors which are responsible for the carcinogenesis process [16-17]. Race, ethnicity, sex, age, inflammatory bowel disease (IBD) [18], and family history are among the non-modifiable factors [19]. On the other hand, modifiable behavioral risk factors refer to lifestyle including diet, physical inactivity, alcohol consumption, and particularly obesity [20], and tobacco use [21-22]. Genetic predisposition is one of the most explored factors and family history seems to have a part in the development of CRC [19, 23]. Thanks to Next Generation Sequencing (NGS)-based analyses several association studies of CRC have identified cancer predisposition genes; most factors causing heritability are still misunderstood and need further study. Currently, the use of multigene panels with NGS strategy is

increasingly used; this is useful to understand the existence or possible expression of this type of pathology, based on the presence of germline mutation, even if often only in terms of risk, or for therapeutic measures or periodic monitoring. Indeed, genetics has a key role in CRC predisposition and in its initiation and progression.

CRC is traditionally divided into sporadic forms (70-80% of the cases) and familial forms that account for 20-30% of the CRC. As many as, 3-5% of familial cases are associated with highly penetrant inherited mutations, and clinical presentations that have been well characterized. In these forms, the understanding of the effect of germline mutations presence, in highly susceptible genes, for the development of colorectal cancer, is important for identifying individuals at risk, for improving cancer surveillance and prevention, and for developing new therapeutic approaches. Hereditary colorectal cancer syndromes can be subdivided as non-polyposis (Lynch syndrome and familial colorectal cancer) and polyposis syndromes. In the hereditary forms, the alteration of a single gene, transmitted through the germline, causes a marked familial predisposition to the development of CRC (Figure 5).

Lynch syndrome (LS), previously known as Hereditary Non-polyposis Colorectal Cancer (HNPCC), is characterized by adenomatous polyps with a high potential for neoplastic degeneration, it is the most common form of hereditary CRC (5-10% of all colorectal cancers) [6,17,23]. Its inheritance is autosomal dominant (AD) with high penetrance and is associated with pathogenic germline variants in DNA mismatch repair (MMR) genes: *MSH2, MSH6, MLH1* and *PMS2;* moreover, several patients with LS show mutations in other genes like *EPCAM, PALB2, POLE, POLD1* [6, 24-25]. Mutations in genes described above predispose to the

development of neoplasms with distinctive molecular phenotypes: DNA mismatch repair deficient (MMR-d) or proficient (MMR-p). MMR-d tumors show a high instability of DNA microsatellites (MSI) and loss of expression of the corresponding DNA mismatch repair protein by immunohistochemistry; it is a hallmark of Lynch Syndrome. Microsatellite instability in tumor DNA is defined as the presence of alternate sized repetitive DNA sequences that are not present in the corresponding germline DNA.

However, MSI is not specific for Lynch syndrome, and approximately 15% of sporadic colorectal cancers also show MSI and its status has prognostic and therapeutic implications.

Because of the accelerated adenoma–carcinoma pathway, patients with LS are advised to undergo frequent, 1–2 yearly, colonoscopy from age 20 years to 25 years for all their life. Besides that, these patients are also at an increased risk for endometrial cancer and other malignancies (e.g., cancers of the small bowel, stomach, ovaries, renal pelvis, ureter, and hepatobiliary system) [26].

The polyposis syndromes are more easily recognized as the physician is alerted by the large number of polyps. The most important is Familial Adenomatous Polyposis (FAP) and it is due to mutations in the *APC* gene, a tumor suppressor in the regulation of WNT signaling. Its inheritance is AD, approximately 30% of affected individuals have no family history and show *de novo* mutations [26] and it is characterized by the appearance at a young age of adenomatous polyps (>100) located throughout the colon. In subjects with overt FAP the risk of developing CRC before age 40 is about 99%; the treatment is surgical and involves prophylactic total colectomy [27].

MUTYH-associated polyposis (MAP) is an autosomal recessive syndrome associated with biallelic germline variants in the base excision repair gene *MUTYH*. Subjects with this type of disease can exhibit a wide range of phenotypes including classic and attenuated polyposis. In the literature have been found monoallelic MUTYH variants associated with a moderate or increased risk for CRC, especially among individuals which are first-degree relative with CRC subjects [6, 28].

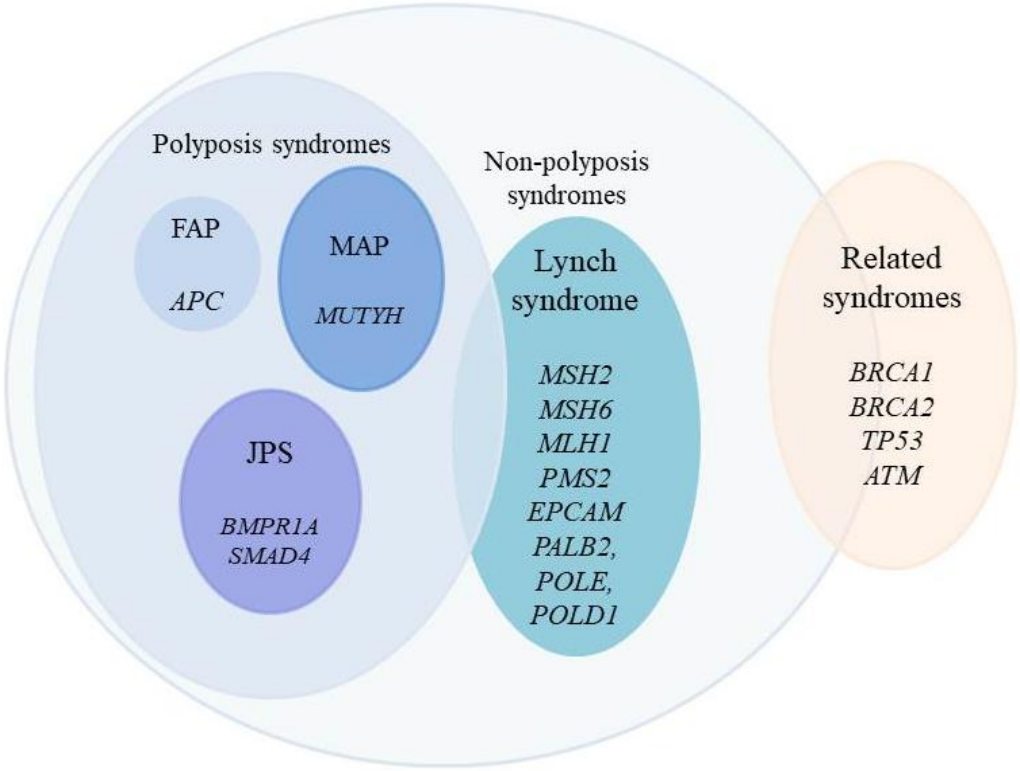


Figure 5. Genes involved in the evolution of CRC. The syndromes related to the CRC can be distinguished into Polyposis and non-polyposis ones. Moreover, there are other genes known to be involved in the progression of CRC (right part of the figure).

Juvenile polyposis syndrome (JPS) is another syndrome characterized by multiple gastric and/or colonic hamartomas. The germline variants are identified in *BMPR1A* and *SMAD4* genes in 50–70% of affected subjects. JPS is associated with high risk of developing colorectal and/or gastric tumor [6,29].

In addition, to all the hereditary diseases occurring above, germline mutations in additional high and moderate penetrance cancer genes have been also associated with increased risk for colorectal neoplasia [30] (see Figure 5).

1.5 Clinical signs and screening methods of CRC

The clinical presentation of CRC often depends on the location, size, and stage of the tumor; however, the most common signs and symptoms include: a persistent change in bowel habits, including diarrhea or constipation or a change in stool consistency, rectal bleeding or blood in the stool, persistent abdominal discomfort, such as cramps, gas or pain, the feeling that your intestines are not emptying completely, weakness or tiredness and unexplained weight loss [16].

Unfortunately, some colorectal cancers can be present without any signs or symptoms. Since CRC does not usually cause symptoms, until the disease is advanced, it is important for people undergo regular screening to find polyps before they become cancerous. The goal of screening is to reduce the number of people who die from cancer; getting screened, and treated early if cancer is founds, reduces the risk of dying from CRC.

The "ideal" screening test should be noninvasive, have high sensitivity and specificity, be safe, readily available, convenient, and inexpensive. For CRC screening, there are multiple approved tests and strategies, each with its strengths and weaknesses. Currently, the guidelines recommend that adults age 45 to 75 be screened for colorectal cancer. The decision to be screened after age 75 should be made on an individual basis [31-32].

Instead, people should start colorectal cancer screening first and/or undergo more frequent screening, if they have any of the following risk factors for colorectal cancer [16]:

- personal history of colorectal cancer or adenomatous polyps;
- strong family history of colorectal cancer or polyps, such as cancer or polyps in a first-degree relative younger than 60 or in 2 first-degree relatives of any age;
- personal history of chronic IBD;
- family history of any hereditary colorectal cancer syndrome, such as FAP, Lynch syndrome, or others.

Currently, there are two level for the screening of CRC; the first level tests provide the Fecal occult blood test (FOBT) or Fecal immunochemical test (FIT) and the recto sigmoidoscopy; if there is a positivity to one of these tests, the screening programs provide for the execution of a colonoscopy as an in-depth examination [31-32]. The second level test involves colonoscopy which allows to examine the entire colon and, if present, to remove any polyps. Up today, no country uses population screening based on colonoscopy as a first level test, given the poor acceptability of the examination and the rate of adverse events found to be between 0.1-0.6%.

In this scenario, genetic tests are becoming increasingly frequent for the diagnosis of hereditary forms of CRC. Using these tests, it is possible to characterize the affected subjects at the molecular level and carry out the predictive test for asymptomatic subjects into families at risk, assess the risk for descendants and make the differential diagnosis between the various syndromes of hereditary colorectal cancers, when necessary. In this way, a correct follow-up can be carried out for

subjects positive to genetic tests; instead, negative subjects could be excluded from the clinical screening program as they present a risk of developing CRC equal to that estimated for the general population. To date, a good methodology for molecular screening of colorectal cancer is the use of multigene panels, using the NGS strategy, for the identification of point germline mutations: this will give a score for identifying predisposition to the disease, to make a more accurate and frequent monitoring by other procedures, i.e.: particularly, colonoscopy with biopsy of uninspected sites.

1.6 Diagnosis, staging and treatment

The diagnosis of CRC is based on an initial evaluation of signs and symptoms and above all on a biopsy taken during a colonoscopy which must confirm the presence of an adenocarcinoma. Therefore, the diagnosis is performed histologically [32-33].

However, some patients can have an incomplete colonoscopy due to colon stenosis or other technical difficulties. In those cases, further computed tomography colonography (CTC) can contribute to an appropriate CRC diagnosis. The most common site of metastases for CRC is the liver, even if the CRC may also spread to the lungs, bones, brain, or spinal cord; for this reason, these sites must be investigated with instrumental imaging examinations [16].

Actually, knowing the stage of the disease at the time of the diagnosis helps the clinical advice for choosing which type of treatment is the best and can help predict a patient's prognosis. There are different stage descriptions for different types of cancer, CRC staging uses the American Joint Committee on Cancer (AJCC) tumor/node/metastasis (TNM) classification [16,34]. This classification is based on

the characteristics of the primary tumor (T stage), involvement of lymph node (N stage) and presence of distant metastasis (M stage). The information of those categories is combined into an overall stage definition (stage I, II, III or IV). Using this CRC staging system, a higher number indicates increasing severity of the cancer (Figure 6).

The CRCs can be also classified histologically based on the degree of preservation of normal glandular architecture and cytological features (well differentiated G1, moderately differentiated G2 or poorly differentiated G3) [35]. CRC can present in 10-20% of the cases a mucinous phenotype; it is characterized by the presence of intracellular accumulation of mucins. This cancer characteristic turns out normally in a very aggressive cancer with a poor prognosis.

For the treatment of CRC, it is necessary a multidisciplinary team, generally including a surgeon, an oncologist, a radiotherapist, and a gastroenterologist. Treatment options and recommendations depend on a several factors, including the type and stage of cancer, possible side effects and the patient's preferences and overall health. For many CRCs, the goal of treatment is to cure the cancer. If cure isn't possible, treatment may be used to shrink the cancer or keep it under control as long as possible.

Surgery is the most common treatment for this disease, depending on several features of the tumor, such as its location and presence of metastasis. If cancer is found only in a single polyp, it can be removed during colonoscopy. In early stages of cancer, treatment can include chemotherapy, radiotherapy, radiofrequency ablation, cryosurgery, or targeted therapy. However, chemo and radio therapy are critical for neo-adjuvant and adjuvant treatment strategies.

A

T stage	
Tx	No information about local tumour infiltration available
Tis	Tumour restricted to mucosa, no infiltration of lamina muscularis mucosae
T1	Infiltration through lamina muscularis mucosae into submucosa, no infiltration of lamina muscularis propria
T2	Infiltration into, but not beyond, lamina muscularis propria
T3	Infiltration into subserosa or non-peritonealised pericolic or perirectal tissue, or both; no infiltration of serosa or neighbouring organs
T4a	Infiltration of the serosa
T4b	Infiltration of neighbouring tissues or organs
N stage	
Nx	No information about lymph node involvement available
N0	No lymph node involvement
N1a	Cancer cells detectable in 1 regional lymph node
N1b	Cancer cells detectable in 2–3 regional lymph nodes
N1c	Tumour satellites in subserosa or pericolicor perirectal fat tissue, regional lymph nodes not involved
N2a	Cancer cells detectable in 4–6 regional lymph nodes
N2b	Cancer cells detectable in 7 or greater regional lymph nodes
M stage	
Mx	No information about distant metastases available
M0	No distant metastases detectable
M1a	Metastasis to 1 distant organ or distant lymph nodes
M1b	Metastasis to more than 1 distant organ or set of distant lymph nodes or peritoneal metastasis

B

	T	N	M
Stage 0	Tis	N0	M0
Stage I	T1/T2	N0	M0
Stage II	T3/T4	N0	M0
IIA	T3	N0	M0
IIB	T4a	N0	M0
IIC	T4b	N0	M0
Stage III	Any	N+	M0
IIIA	T1–T2	N1	M0
	T1	N2a	M0
IIIB	T3–T4a	N1	M0
	T2–T3	N2a	M0
	T1–T2	N2b	M0
IIIC	T4a	N2a	M0
	T3–T4a	N2b	M0
	T4b	N1–N2	M0
Stage IV	Any	Any	M+
IVA	Any	Any	M1a
IVB	Any	Any	M1b

Figure 6. Staging of CRCs. A) Classification of colorectal cancers according to T stage (invasion depth), N stage (lymph node involvement) and M stage (presence of metastasis). B) Overall CRC staging from early stages (Stage I) to advanced stages (Stage IV). Adapted from "Colorectal cancer" (2014) [16].

1.7 Molecular pathogenesis of CRC

Colorectal cancer is a multifactorial disease and, one of the central aspects of its formations is the accumulation of acquired genetic and/or epigenetic changes. These alterations are responsible of the activation of oncogenes or inactivation of oncosuppressor genes that transform normal glandular epithelial cells trough early neoplastic lesions into invasive adenocarcinomas [35-37].

The neoplastic transformation is characterized by two morphological pathways: the Conventional and alternative adenoma-carcinoma (Figure 7). In the first case, the so-called adenoma-carcinoma is histologically homogenous, and its progression is also the consequence of molecular mechanisms: chromosomal instability (CIN) and/or microsatellite instability (MSI) [16,29,36]. The most common form of genomic instability is CIN, which is found in 85% of CRCs. CIN is recognized by the presence of numerical chromosome changes or multiple structural aberrations of the chromosomes. It is characterized by mutations in tumor suppressor genes and oncogenes, such as *APC, TP53* and *KRAS* [38]; these bring to a consequent dysregulation of the Wnt/β-catenin, MAPK, PI3K and TGF-β signalling pathways.

Alternatively, MSI occurs in approximately 15-20 % of all CRC cases. MSI is due to an accumulation of insertions/deletions of short nucleotide repeats (microsatellites) consecutive to alteration in genes encode proteins that repair mismatched DNA bases [34]. If this functionality is lost due to gene mutation, an adenoma can form and may, within a few years, progress to malignancy as well. The genes involved in this process are *MLH1, MSH2, MSH6,* and *PMS2* [36-39]. Depending on the number of the microsatellites found, tumors have been classified into: (i) high, “MSI”, (ii) low, “MSI-L” or (iii) stable, “MSS”.

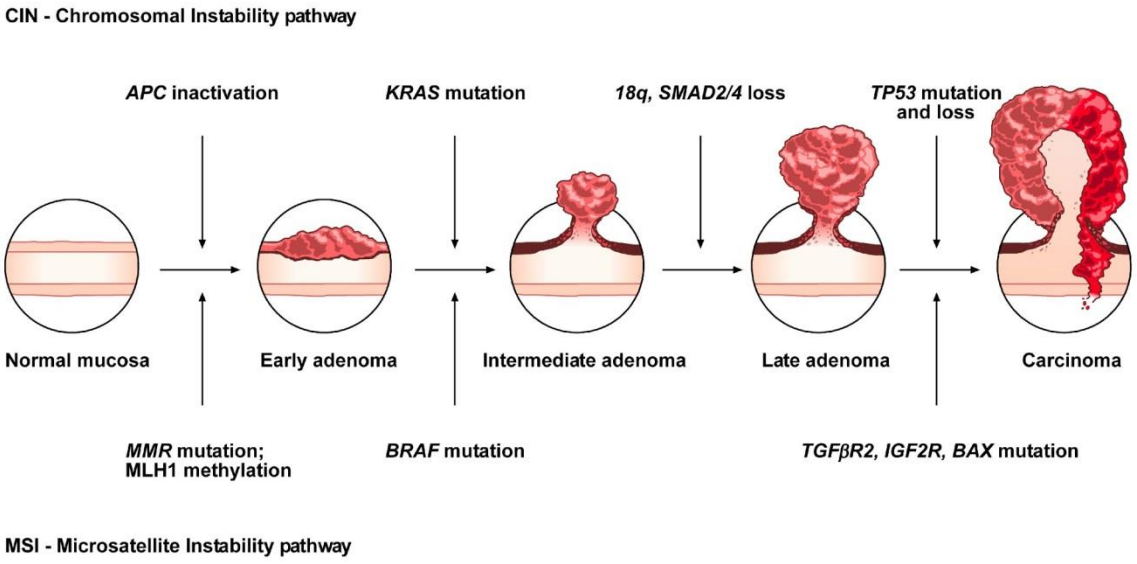


Figure 7. Conventional adenoma-carcinoma pathway. CIN is characterized first by a mutation in *APC* gene followed by mutations in the *KRAS* and *SMAD4* genes. MSI is due to mutations in genes involved in the mismatch repair such as *MLH1*. Then, mutation in *BRAF* participates in the progression to the intermediate and late stages of carcinogenesis. Adapted from "The Molecular Hallmarks of the Serrated Pathway in Colorectal Cancer" [36].

Instead, regarding the alternative pathway, it is characterized histologically by sessile serrated adenomas/polyps as precursor lesions rather than tubular adenoma. These lesions, at the molecular level, are characterized by mutations in *APC* and *BRAF* genes; moreover, CpG island methylator phenotype (CIMP) mechanism is involved. The 20-30% of CRCs carries CIMP phenotype. CIMP mechanism involves epigenetic hypermethylation of CpG islands in promoter regions of tumor suppressor genes, preventing them from undergoing transcription and causing an inactivation of themselves [39, 40].

The described pathways provide a basic outline of the molecular mechanisms involved in the CRC development, but further studies and insights are needed.

1.8 New strategy for precision medicine: 3D culture

In the last decades, knowledge on the origin of tumors and relative mechanisms involved has increased extremely. Despite great advances in the treatment of cancer patients, cancer continues to be the second leading cause of cancer-related death worldwide [2,3].

Unfortunately, the creation of a model that resembles the patient's tumor is still a problem; therefore, many drugs tested *in vitro* fail during clinical trials [41,42].

For many decades the 2D *in vitro* cultures have been used but cannot reproduce the real complexity of the anatomical structures existing in human body. Over the years, primary patient-derived tumor xenografts (PDTXs) have been widely used for functional genomic studies in a variety of tumor types [42]. PDTXs are generated transplanting freshly derived patient material subcutaneously or, orthotopically into immunodeficient mice. The limits of this model are linked to the use of animals and to the success of the engraftment.

In the last years, patient derived organoids (PDOs) have been developed. This technology is based on the ability of a group of cells growing in a three-dimensional (3D) matrix, usually the matrigel, using tissue-specific growth factors. Organoids grow starting from pluripotent stem cells (PSCs) and from adult stem cells (ASCs) by mimicking human development or organ regeneration *in vitro* [43,45] (Figure 8). PSC-stabilized organoids form structures that are found during embryonic development. The process of forming these organoids is multi-step and takes many months; this type of culture is mainly used for studying development, infectious and genetic diseases. In contrast to PSC-derived organoids, ASC-derived epithelial

organoids recapitulate adult tissues, and they can be established only from tissue compartments with regenerative capacity [42,45].

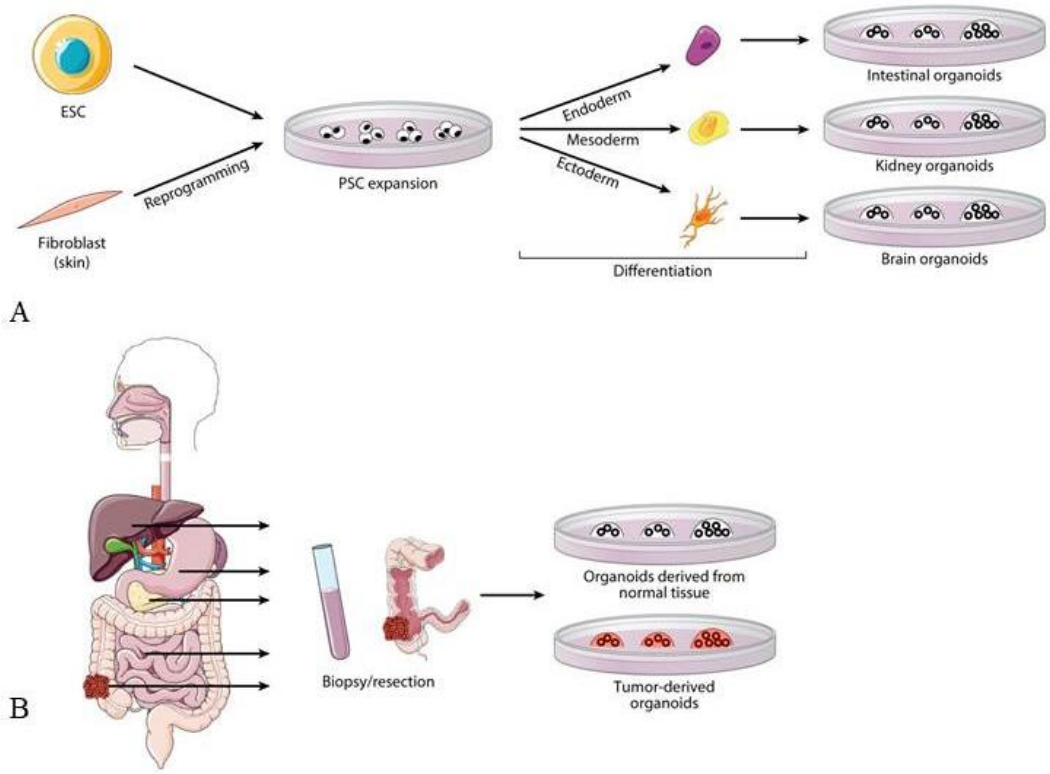


Figure 8. Different strategies for Organoid formation. A) Pluripotent stem cell (PSC)–derived organoids use induced PSCs derived from cells like skin fibroblasts or embryonic stem cells (ESCs); then, differentiated into three germ layers (endoderm, mesoderm, and ectoderm). Different protocols are used to obtain the tissue of interest. (B) Adult stem cell–derived organoids use tissues from biopsies or resections from many organs; then the cells are used to obtain, respectively, organoid cultures derived from normal epithelial tissue and tumor tissue–derived organoid cultures. Adapted from "Human Organoids: Tools for Understanding Biology and Treating Diseases" (2020) [42].

They can derive from both healthy and tumoral tissues and the great advantage is the presence of many of the cell types present in the organ from which they derived; moreover, PDOs can also recapitulate some of the spatial organization of the primary tissue [40,43]. The process of forming these organoids is multi-step, it takes about 7 days for the formation of the structures, and they can be expanded long term by keeping the genetical stability. Thanks to this ability, they are heavily used for personalized therapies.

The first type of stabilized organoid from ASC was from the gut-intestine after the identification of Lgr5 as a marker of Wnt-driven adult gut stem cells [47]. During the years, modifying the cell isolation procedures and some growth factors, different organoids were developed from other organs, as colon [48], prostate [49], gastric [50], breast [51], pancreas [52], liver [53], lung [54], kidney [55] and brain cancers [56].

The organoid technology is versatile: cultures can be started from small tissue sections obtained by biopsy or surgical specimens. The tissue fragments are enzymatically disrupted, seeded into a multi-well and during the time they form cystic structures lined by a polarized epithelium.

Obviously, there are hurdles and limitations associated with using organoids. First, organoid culture requires the use of an animal-based matrix extract; moreover, they may carry unknown pathogens and are potentially immunogenic when transplanted to human and it can limit the use of organoids in a clinical transplantation setting [42,43].

In summary, thanks to this new technology, the correlations between the clinical characteristics of the patient and the pre- and post-treatment analyses can be useful for solving technical challenges and for implementing successful applications in the clinic.

2. Background and aim of the study

Cancer is one of the leading causes of death worldwide; in 2020, the most common cancers diagnosed were: breast cancer with 2.26 million cases, lung cancer with 2.21 million cases, colon and rectum cancers with 1.93 million cases. Among common malignancies, CRC accounts for one of the largest numbers of familial cases.

The accumulation of genetic anomalies (SNV, small INDELs and/or CNVs and chromosomal alterations) and epigenetic modifications in genes that control proliferation, differentiation, death and integrity of the genetic pattern plays an important role in the development of CRC. The mutational state in predisposing genes is a potential prevention tool. Thus, high risk subjects can be identified based on their state and be considered candidates for targeted prevention and/or therapeutical strategies. In addition to germline mutations, which in many cases are also driver mutations, somatic mutations can help in making treatment decisions.

In recent years, the use of 3D *in vitro* cultures has emerged for their ability to mimic some features and possibly the fate of solid tumors. Consequently, we established PDOs from CRC patients and compared four genomes sequences derived from each patient (blood, PDOs, tumor-derived and paired-healthy tissue). To this end, we customized a multi-gene panel using NGS strategy to identify germline mutations in genes associated with CRC. Moreover, we used this multi-gene panel to identify somatic mutations in 3 genomes (PDOs, tumor-derived and paired-healthy tissues). These somatic mutations can help to shed light on tumoral cells present in PDOs in terms of drug screening and cancer treatment (Figure 9).

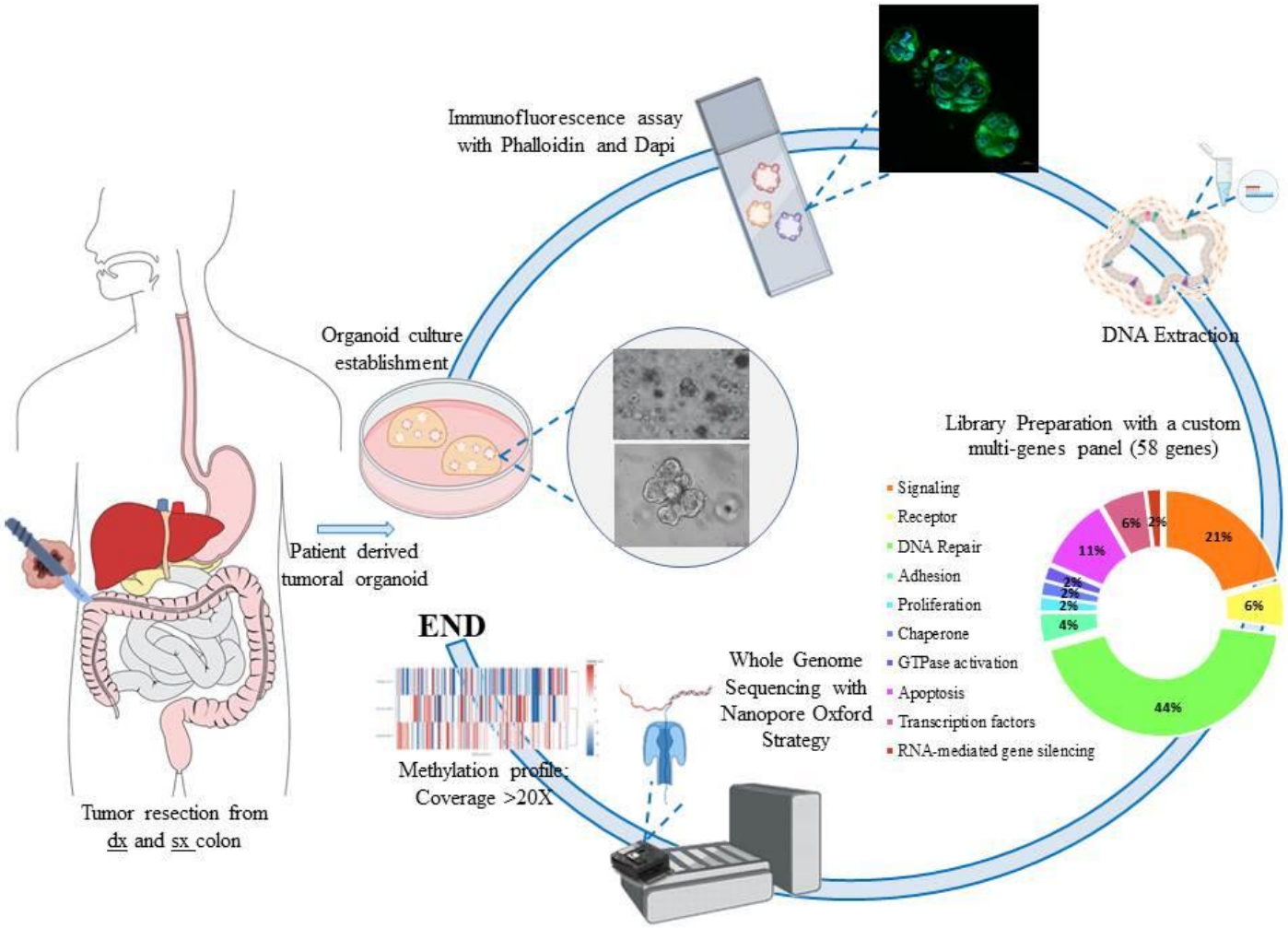


Figure 9. Graphic description of the various phases of the experiments performed. Starting from the left part of the image all the steps are shown.

3. Materials and Methods

3.1 Patients' enrollment

Eighty patients were enrolled for this study and all of them were affected by colorectal cancer (CRC) and underwent to surgery. All patients were followed up at the University of Naples Federico II: among these, the patients were selected from U.O.C. of General Surgery and Mininvasive oncology, and from DAI of Gastroenterology, Endocrinology and Endoscopic Surgery; all research investigations at molecular and cellular level were carried out at the CEINGE-Biotecnologie Avanzate in Naples. Informed written consent was obtained from all the patients enrolled in this study according to the procedure established by the Second Helsinki Declaration and according to the Italian and Local Regulations (Ethics Committee number 318/20). In our cohort of patients there are 36 female and 43 male and the mean age of the patients at the time of the diagnosis and the surgery is 70 and 71 years, respectively, and there is no difference between the two sexes (for one patient detailed information was missed). All patients show colon adenocarcinoma with different anatomical localization: 25 ascending colon (right-sided), 5 transverse colon, 4 splenic flexure, 19 descending colon (left-sided), 14 sigmoid colon, 5 junction sigma-rectum and 7 rectum.

Thirty-six patients had familiarity with oncology diseases and report having at least one relative with a tumor, while 39 have no familiarity with oncology diseases and for 4 this information is missing (see table 1).

For this project, we designed a specific scheme for the collection of biological materials: peripheral blood was collected for each patient at the time of admission in

the hospital. At the time of the surgery 100 mg of neoplastic mucosa and the same amount of adjacent healthy mucosa were taken; in addition, a full-thickness piece of tumoral tissue of about 300 mg was biopsied, collected in Advanced Dulbecco's modified Eagle medium/ F12+++ [as supplemented with 1X of Primocin -Invivogen, CA, USA, 1X Hepes (1M) and 1X Glutamax- Thermo Fisher Scientific, Waltham, MA, USA] (Figure 10).

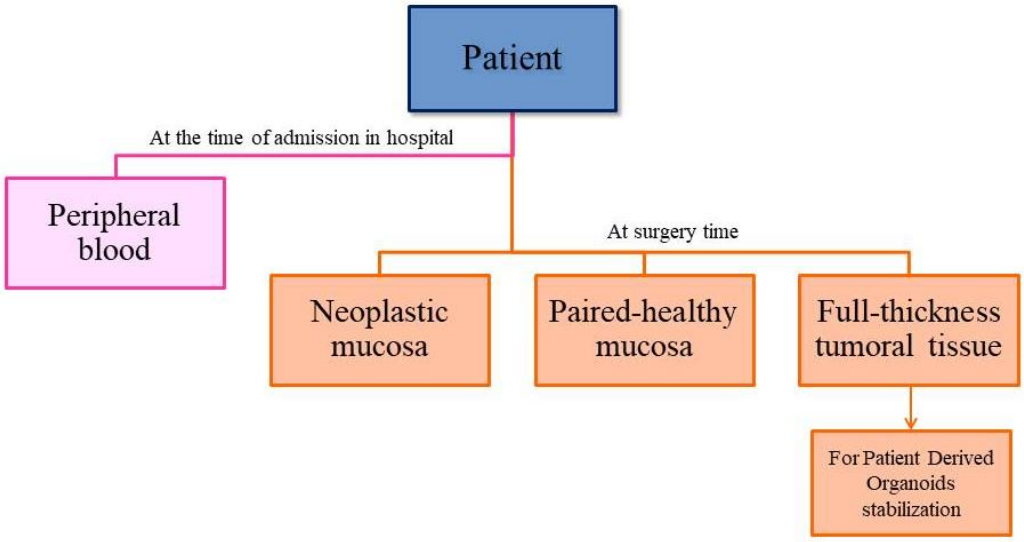


Figure 10: Scheme of the sample collection. From the same patient we obtained different samples: at the time of admission in hospital and at the time of the surgery.

Table 1: Clinical and Histological information of 80 enrolled patients. On the left part of the table are shown all the clinical information obtained from each patient, on the right part are reported all the information transmitted by the two surgery divisions after the routinary histological examination. "-"indicates that these data are not available.

Clinical information									Histological Information					
N	ID Patient	Male Female	Age	Age at the diagnosis	Localization of cancer	Obesity (YES/NO)	Familiarity with oncological diseases	Other tumor	Histotype	T (Tumor)	N (Node)	M (Metastasis)	G (Grade)	Lymph node analyzed (n°)
1	CR_01	M	66	65	Splenic flexure	NO	NO	NO			-			
2	CR_02	F	87	86	Ascending colon	NO	NO	NO			-			
3	CR_03	M	55	54	Descending colon	NO	NO	NO			-			
4	CR_04	M	69	68	Descending colon	NO	YES	NO			-			
5	CR_05	F	80	79	Sigmoid colon	NO	YES	NO			-			
6	CR_06	M	71	70	Sigmoid colon	NO	YES	K Prostate, Liver metastases, Bladder metastases			-			
7	CR_07	F	92	91	Rectum	NO	YES	NO			-			
8	CR_09	F	71	70	Transverse colon	NO	YES	NO			-			
9	CR_10	F	82	81	Rectum	YES	NO	NO	Micinous type adenocarcinoma	T3	N0	-	G4	8
10	CR_11				-						-			
11	CR_13	F	65	63	Descending colon	NO	NO	NO			-			
12	CR_14	M			Ascending colon	YES	NO	NO			-			
13	CR_15	F	84	83	Ascending colon	YES	YES	NO	Infiltrating adenocarcinoma	T1	N0	-	G1	34
14	CR_16	F	91	90	Sigmoid colon	NO	NO	NO	Poorly differentiated adenocarcinoma	T3	N0	Mx	G3	12
15	CR_17	M	65	64	Rectum	NO	NO	NO			-			
16	CR_18	M	70	69	Rectum	NO	NO	NO			-			
17	CR_19	F	78	77	Ascending colon	YES	YES	NO			-			

Clinical information									Histological Information					
N	ID Patient	Male Female	Age	Age at the diagnosis	Localization of cancer	Obesity (YES/NO)	Familiarity with oncological diseases	Other tumor	Histotype	T (Tumor)	N (Node)	M (Metastasis)	G (Grade)	Lymph node analyzed (n°)
18	CR_20	F	89	88	Ascending colon	NO	NO	NO	Moderately differentiated adenocarcinoma	T3	N1b	-	G2	28
19	CR_21	M	70	69	Sigmoid colon	YES	YES	NO			-			
20	CR_22	F	80	79	Ascending colon	NO	NO	K rectum			-			
21	CR_23	F	72	71	Transverse colon	NO	YES	NO			-			
22	CR_24	F	26	25	Rectum	NO	NO	NO			-			
23	CR_26	F	78	77	Sigmoid colon	NO	NO	NO	Colic adenocarcinoma	T3	N0	-	G3	12
24	CR_27	M	80	79	Ascending colon	YES	YES	SI			-			
25	CR_29	M	66	65	Ascending colon	NO	YES	NH lymphoma			-			
26	CR_30	M	67	66	Ascending colon	NO	NO	NO			-			
27	CR_32	F	83	82	Transverse colon	NO	-	NO			-			
28	CR_33	F	78	78	Rectum	YES	YES	NO	Poorly differentiated adenocarcinoma	T3	N0		G2	20
29	CR_34	M	58	58	Sigmoid colon	YES	NO	Oesophageal adenocarcinoma	Adenocarcinoma	T2	N1		G1	19
30	CR_35	M	53	53	Descending colon	YES	NO	NO	Moderately differentiated adenocarcinoma	T3	N0		G2	20
31	CR_36	F	73	73	Descending colon	YES	YES	NO			-			
32	CR_37	M	71	71	Sigmoid colon	YES	NO	NO	Poorly differentiated adenocarcinoma	T2	N0	Mx	G3	21
33	CR_38	F	75	75	Sigmoid colon	NO	NO	NO	Well-differentiated adenocarcinoma infiltrating the perivisceral adipose tissue	T3	N0	-	G1	12

Clinical information									Histological Information					
N	ID Patient	Male Female	Age	Age at the diagnosis	Localization of cancer	Obesity (YES/NO)	Familiarity with oncological diseases	Other tumor	Histotype	T (Tumor)	N (Node)	M (Metastasis)	G (Grade)	Lymph node analyzed (n°)
34	CR_39	M	82	82	Ascending colon	NO	NO	K lung	Slightly differentiated colonic carcinoma, infiltrating the external serous margin	T4	N1	-	G3	16
35	CO_01	M	88	87	Ascending colon	NO	NO	NO	Micinous type adenocarcinoma	T2	N0	M0	G1	12
36	CO_02	M	72	71	Ascending colon	NO	YES	NO	Micinous type adenocarcinoma	T3	N0		G1	17
37	CO_03	F	64	63	Descending colon	NO	YES	Breast cancer	Micinous type adenocarcinoma	T3	N0	M1	G2	38
38	CO_04	F	59	58	Sigmoid colon	NO	YES	NO	Adenocarcinoma	T2	N0			21
39	CO_05	M	71	70	Ascending colon	NO	YES	NO	Micinous type adenocarcinoma	T3	N0	M0	G1	26
40	CO_06	F	78	77	Ascending colon	NO	NO	NO	Adenocarcinoma	T2	N0	M0	G2	
41	CO_07	F	75	74	Ascending colon	YES	NO	Gastrointestinal polyps	Moderately differentiated adenocarcinoma	T3	N0		G2	12
42	CO_08	M	88	87	Ascending colon	NO	YES	Liver metastases	Moderately differentiated adenocarcinoma	T4	N1	M1	G2	16
43	CO_09	F	61	60	Ascending colon	NO	YES	NO	Adenocarcinoma	T2	N0	M0	G2	13
44	CO_10	M	83	82	Ascending colon	YES	YES	NO	Adenocarcinoma	T3	N2b	M0	G2	22
45	CO_11	M	80	79	Rectosigmoid junction	NO	NO	NO	Adenocarcinoma	T4	N1	M0		-
46	CO_12	M	84	83	Sigmoid colon	YES	NO	NO	Moderately differentiated adenocarcinoma	T3	N1a		G2	12

Clinical information									Histological Information					
N	ID Patient	Male Female	Age	Age at the diagnosis	Localization of cancer	Obesity (YES/NO)	Familiarity with oncological diseases	Other tumor	Histotype	T (Tumor)	N (Node)	M (Metastasis)	G (Grade)	Lymph node analyzed (n°)
47	CO_13	F	55	54	Transverse colon	NO	YES	NO				-		
48	CO_15	M	72	71	Ascending colon	NO	NO	NO	Adenocarcinoma	T3	N1	M0		-
49	CO_16	M	71	70	Ascending colon	YES	YES	Previous bladder polyps	Micinous type adenocarcinoma	T3	N0	Mx	G1	18
50	CO_17	F	59	58	Sigmoid colon	YES	YES	NO	Well-differentiated adenocarcinoma	T3	N1	-	G1	22
51	CO_18	M	84	83	Ascending colon	NO	NO	NO	Moderately differentiated adenocarcinoma	T2	N0	-	G2	28
52	CO_19	M	61	60	Splenic flexure	NO	YES	NO	Tubulo-villous adenoma			-		4
53	CO_20	M	65	64	Sigmoid colon	NO	YES	NO	Moderately differentiated adenocarcinoma	T3	N0	-	G2	13
54	CO_21	M	77	76	Descending colon	NO	YES	Atrial myxoma	Mucinous colonic adenocarcinoma	T4b	N0	-	G3	31
55	CO_22	M	69	68	Sigmoid colon	YES	YES	K lung, not small cell	Moderately differentiated adenocarcinoma	T3	N0	-	G2	-
56	CO_23	F	84	83	Descending colon	YES	YES	NO	Moderately differentiated adenocarcinoma	T3	N0	-	G2	31
57	CO_24	M	77	78	Descending colon	YES	YES	NO	Moderately differentiated adenocarcinoma	T1	Nx	-	G2	18
58	CO_25	F	65	65	Descending colon	NO	YES	NO	Multiple diverticulosis			-		6
59	CO_26	M	73	73	Descending colon	NO	NO	K prostate	Moderately differentiated adenocarcinoma	T2	N0	-	G1	32

Clinical information									Histological Information					
N	ID Patient	Male Female	Age	Age at the diagnosis	Localization of cancer	Obesity (YES/NO)	Familiarity with oncological diseases	Other tumor	Histotype	T (Tumor)	N (Node)	M (Metastasis)	G (Grade)	Lymph node analyzed (n°)
60	CO_27	F	78	78	Transverse colon	YES	YES	NO	Moderately differentiated adenocarcinoma	T3	N1	M1	G1	6
61	CO_28	F	57	57	Sigmoid colon	YES	NO	NO	Moderately differentiated adenocarcinoma	T3	N1	-	G1	36
62	CO_29	M	49	49	Rectosigmoid junction	NO	NO	NO	Moderately differentiated adenocarcinoma	T3	N2	-	G2	28
63	CO_30	M	62	62	Descending colon	NO	YES	NO	Well-differentiated adenocarcinoma of the sigmoid rectum	T3	N0	Mx	G2	23
64	CO_31	F	64	64	Descending colon	NO	YES	NO	Well-differentiated adenocarcinoma	T3	N0	Mx	G1	26
65	CO_32	F	61	61	Descending colon	NO	NO	NO	Well-differentiated adenocarcinoma	T4a	N2b	Mx	G2	32
66	CO_33	M	87	87	Rectosigmoid junction	NO	NO	NO	Well-differentiated adenocarcinoma	T4a	N1b	Mx	G3	34
67	CO_34	M	61	61	Descending colon	NO	NO	Previous ascending colon	Poorly differentiated colon adenocarcinoma	T2	N1b	Mx	G3	15
68	CO_35	M	51	50	Descending colon	NO	YES	YES	Moderately differentiated adenocarcinoma	T3	N0	Mx	G2	40
69	CO_36	F	73	73	Descending colon	NO	NO	NO	Moderately differentiated adenocarcinoma	T2	N0	-	G2	23
70	CO_37	F	61	61	Descending colon	YES	YES	NO	Moderately differentiated adenocarcinoma	T3	N1a	-	G2	35

Clinical information									Histological Information					
N	ID Patient	Male Female	Age	Age at the diagnosis	Localization of cancer	Obesity (YES/NO)	Familiarity with oncological diseases	Other tumor	Histotype	T (Tumor)	N (Node)	M (Metastasis)	G (Grade)	Lymph node analyzed (n°)
71	CO_38	F	43	43	Descending colon	NO	NO	NO	Poorly differentiated colon adenocarcinoma	T3	N0	Mx	G3	36
72	CO_39	M	76	76	Splenic flexure	YES	NO	NO			-			
73	CO_40	F	69	69	Ascending colon	NO	NO	Liver metastases	Poorly differentiated colon adenocarcinoma	T3	N2	M1	G3	13
74	CO_41	M	66	-	Rectosigmoid junction	-	-	-	Well-differentiated adenocarcinoma	T3	N0	Mx	G1	8
75	CO_42	M	87	87	Rectosigmoid junction	YES	-	NO	Well-differentiated adenocarcinoma			-		9
76	CO_43	F	71	71	Ascending colon	YES	NO	NO	Advanced adenoma			-		15
77	CO_44	M	74	74	Splenic flexure	NO	YES	NO	Well-differentiated adenocarcinoma	T3	N0	-	G1	17
78	CO_45	M	68	68	Ascending colon	NO	NO	NO	Well-differentiated adenocarcinoma	T1	N0	-	G1	34
79	CO_46	M	68	68	Ascending colon	YES	NO	NO	Well-differentiated adenocarcinoma	T3	N0	-	G1	26
80	CO_47	M	73	73	Rectum	YES	NO	NO	Poorly differentiated colon adenocarcinoma	T2	N0	Mx	G2	37

3.2 Patient's preparation for surgery

For colon surgery, patients first follow a slag-free diet for approximately seven days. The day before the intervention they do liquid diet. Moreover, they undergo to mechanical bowel preparation with low pressure enemas.

3.3 DNA Extraction from blood, tumoral and paired-healthy tissues

DNA was extracted separately from: peripheral blood samples collected in Ethylene Diamine Tetra Acetic acid (EDTA), and from tumoral and paired-healthy tissues, using the Promega 16 LEV blood DNA purification kit and the Maxwell, instrument Promega Maxwell (Promega Corporation, 2800 Woods Hollow Road Madison, WI, USA), according to manufacturer's instructions.

In particular, 300 µl of Lysis Buffer and 30 µl of Proteinase K were added to 300 µl of blood or to 10 mg of tissue from each study subject. Each tube was vortexed for 10 seconds and incubated in a thermoblock for 20 minutes at 56°C. During this time, the reagent cartridge and the instrument were prepared. The reagent cartridge to be used was placed in the Maxwell 16 Lev Cartridge Rack and, after the incubation, each sample (~650 µl) was transferred into the cartridges. At the end of the extraction protocol, the quality and quantity of extracted genomic DNA (gDNA) were measured using a NanoDrop ND-1000 Spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA).

3.4 Digestion of the tumoral tissue for organoids stabilization

The protocol we followed provides a method for establishing human intestinal Adult Stem Cells (ASC)- derived organoid cultures from fresh biopsies of colon.

Tissue is dissociated into small epithelial fragments that are embedded into extracellular matrix “domes” using Matrigel Matrix (Corning, NY, USA) and then supplemented with medium containing growth factors for organoid expansion (Table 2).

Table 2. Medium for PDOs from Colorectal cancer.

Reagent	Stock Concentration	Final Concentration		F.D	Volume for 50 mL of medium	
A83-01	25 mM	500	nM	1000	50	µl
mEGF	50 µg/mL	50	ng/ml	1000	50	µl
Gastrin	0,021 mg/ml = 10000nM	10	nM	1000	50	µl
mNoggin	100 µg/mL	100	ng/mL	1000	50	µl
IGF-I	100 µg/mL	100	ng/mL	1000	50	µl
N-acetylcysteine	81.5 mg/ml = 500mM	1,25	mM	400	125	µl
Nicotinamide	122 mg/mL= 1000mM	10	mM	100	500	µl
B27 supplement	-	1	µl/mL	-	1000	µl
AdvDMEM/F12+++					48125	µl

Surgically resected intestinal tissues were obtained from 80 patients as described above; the mucosa of the colon was cleaned from any residual faeces with sterile gauzes and was washed with a physiological solution; then, a slide of about 50 mg of full thickness tumoral tissue was biopsied on the opposite side of the lesion (possibly longitudinally along one of the tapeworms) and preserved into Advanced Dmem/F12+++.

More in detail, once the tissue arrived in the laboratory, it was washed before with cold Phosphate buffered saline (PBS)- Thermo Fisher Scientific, Waltham, MA, USA, after, with Primocin, antimicrobial agent for primary cells, and Penicillin-Streptomycin (Thermo Fisher Scientific, Waltham, MA, USA). The tissue was

chopped into approximately 5-mm pieces and then incubated for 45 minutes at 37°C in digestion buffer: Advanced DMEM/F-12 +++, 40 mg of Collagenase II (Thermo Fisher Scientific, Waltham, MA, USA) and 10 mg of Dispase II (Thermo Fisher Scientific, Waltham, MA, USA). At the end of the incubation, the fragments were allowed to settle down under normal gravity for 1 minute on ice and were centrifuged at 150–200g for 5 minutes; the pellet was washed with supplemented Advanced Dmem/F12+++ and centrifuged again as described before.

The supernatant was removed and 2ml of TrypLE (Thermo Fisher Scientific, Waltham, MA, USA) and 20µl of DNase (Sigma-Aldrich, St. Louis, Missouri, USA) were added to the pellet and incubated for 5 minutes at 37°C; cool supplemented medium was added and then centrifuged at 150-200g for 5 minutes. Finally, the supernatant was aspirated and the appropriate amount of Matrigel was added. The cells into the matrigel were plated into 24-multiwell plate (Sartstedt, Nümbrecht, Germany) and, after the matrigel was polymerized for 10 minutes at 37°C in the incubator, a supplemented hot medium was added (see Table 2).

Every 24 hours the cultures were checked, and the culture medium was changed every 2 days.

After 7 days, we expanded the organoids by mechanical disruption and splitting; this process is important for the growth of them (see Figure 11). All the images were taken with DMI4000 B-Leica inverted optical microscope to acquire images of our organoids in bright field

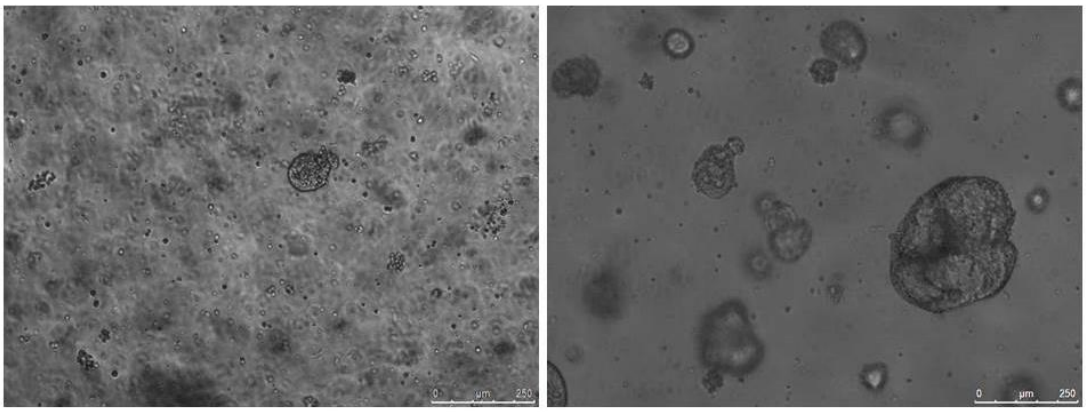


Figure 11. Two different PDOs after 7 days. In the two figures are shown two different PDOs, before the splitting, derived from A) patient CR_32 and B) CR_34.

For the split, the formed organoids were recovered with 10 ml of cold Advanced Dmem/F12+++ and were centrifuged at 150-200 g for 5 minutes. The supernatant was removed, 2 ml of Cell Recovery (Corning, Thermo Fisher Scientific, Waltham, MA, USA) and 8 ml of Advanced Dmem/F12 +++ were added to the pellet, then incubated for 1 hour on ice. At the end of the incubation, the organoids were centrifuged as always. The supernatant was removed, and 2 ml of TripLe were added to dissociate the formed structures; then, the cells were incubated for 2 minutes at 37°C and to stop the reaction 8 ml of medium were added. After the dissociation step, the cultures were centrifuged, the pellet was resuspended into the matrigel and the organoids were plated as described before.

3.5 DNA extraction from tumoral organoid

For the extraction of DNA from Patient Derived Tumor Organoids (PDTOs), we isolated the formed structures contained in 4 domes of matrigel. The process is the same described before in the phase of the split; instead of resuspending the formed

organoids in the matrigel, they were washed with PBS and then DNA was extracted according to the manufacturer's instruction of QIAamp DNA Blood Mini Kit (Qiagen, Hilden, Germania).

As described in the protocol, 200 µl of Buffer AL were added to the sample following incubation step of 10 minutes at 56°C; then, 200 µl of Ethanol 100% were added. The mixture was pipet into a DNeasy Mini spin column that was centrifuged for 1 minute at 6000g. The flow-through was discarded and 500 µl of Buffer AW1 were added. Centrifugation step was performed, 500 µl of Buffer AW2 were added to the mini spin column and the sample was centrifuged twice for 1 minute at 6000g. At the end, the DNA was eluted in 60 µl of Buffer EB.

The quality and quantity of extracted genomic DNA (gDNA) were measured using a NanoDrop ND-1000 Spectrophotometer.

3.6 Multi-gene panel sequencing design

The entire study group was analysed with a custom multi-genes panel, formed by 58 genes, specifically designed in our laboratory. The genes included in the panel have been carefully selected after deep study of the literature. The custom design was realized using the web-based application HaloPlex SureDesign site (www.genomics.agilent.com). Our multi-genes panel accounts for 1032 target regions and contains 25.907 probes for a total size of 564.849 kbp. Our custom panel includes all coding exons for each gene, at least 50 bp at exon boundaries on each side (5' and 3') and also the 5' promoter sequence of each gene and the 3' UTR region. The 58 genes selected are associated to different tumors such as breast and ovarian cancer, colon and prostate cancers. The genes selected in this panel are

responsible for DNA repair, apoptosis, transcription factors and other functions, as shown in the Figure 12.

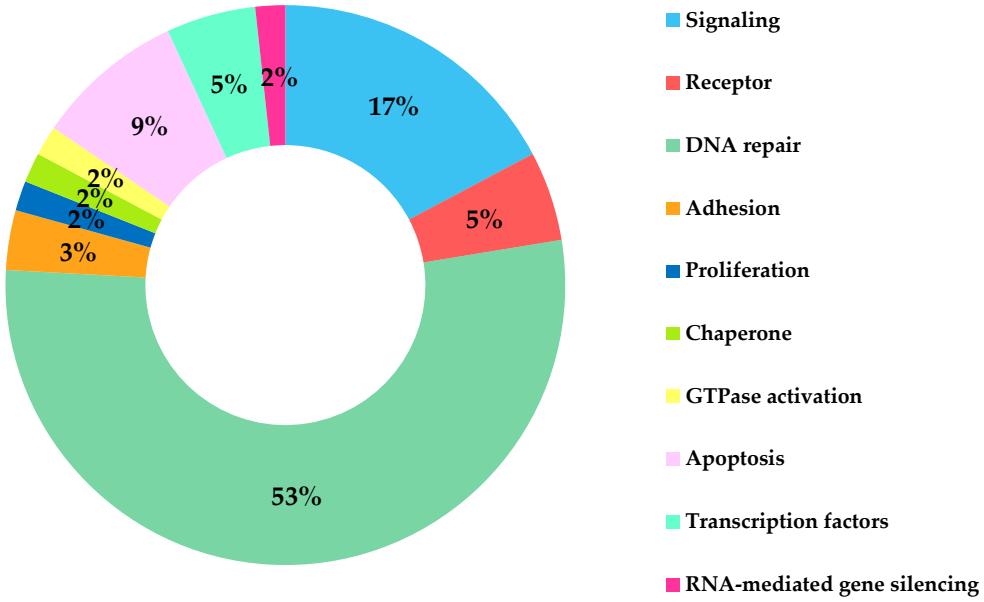


Figure 12. Distribution of the different genes in multi-genes panel. Based on the colours assigned are reported the percentage of each class of genes in our 58 multigene panel.

An enriched DNA library has been obtained using the SureSelect QXT Target Enrichment (Agilent Technologies, Santa Clara, CA, USA), following manufacturer's instructions.

3.7 Library Preparation of multi-genes panel

For each sample to be sequenced, an individual target-enrichment, index library was prepared. The gDNA extracted as described before (from blood, tumoral and healthy tissue and PDTOs) was fragmented and adaptors were added in a single enzymatic step. The adaptor-tagged DNA library was purified and amplified. Next,

each library was hybridized using SureSelect^{QXT} capture library. The resulting libraries were recovered using streptavidin magnetic beads and a post-capture PCR amplification was carried out.

The SureSelect^{QXT} protocol is optimized for the fragmentation of 50 ng gDNA; so, before starting the procedure, all gDNA samples were quantified using the Invitrogen Qubit dsDNA BR assay (Thermo Fisher Scientific, Waltham, MA, USA). To the gDNA prepared before were added 17 µl of Sureselect QXT buffer and 2 µl of Sureselect QXT Enzyme Mix. The samples were incubated into the thermal cycle as follows: 45°C for 10 minutes, 4°C for 1 minute and 4 °C hold. At the end of the incubation, 32 µl of 1x SureSelect QXT Stop Solution were added to each fragmentation reaction and purification step was carried out with AMPure XP beads (Beckman Coulter, Fullerton, CA, USA). To each Adaptor-tagged sample were added 52 µl of the homogeneous beads and the obtained solutions were incubated for 5 minutes at room temperature; then, the tubes were kept on the magnetic plate and the cleared supernatants were discarded. A wash step was performed by adding 200 µl of 70% EtOH into each sample tube, the Ethanol was removed, and the step was repeated for a total of two washes. The tubes were then air-dried with open lids at room temperature until the residual ethanol was evaporated; beads were resuspended with 11 µl of Nuclease-free water. The purified adaptor-tagged gDNA was repaired and 40 µl of PCR Mix for each sample were prepared as follows: 25 µl of Nuclease-free water, 10 µl of Herculase II 5x Reaction Buffer, 0.5 µl of 100mM dNTP Mix, 2.5 µl of DMSO, 1 µl of Sureselct QXT Primer Mix and 1 µl of Herculase II Fusion DNA Polymerase. The samples were incubated in the thermal cycler with the scheme reported below: 68°C for 2 minutes, 38°C for 2 minutes, 98°C for 30 seconds, 57°C

for 30 seconds and 72°C for 1 minutes; the last three steps were repeated for a total of 8 cycles, 72°C for 5 minutes and 4°C hold. At the end of the PCR step, the samples were purified with AMPure XP bead, as described before. At that time, the libraries were checked with D1000 ScreenTape System using 2200 TapeStation (Agilent Technologies, Santa Clara, CA, USA), following manufacturer's instruction to verify the presence of the peak of DNA fragment positioned between 245 to 325 bp (Figure 13).

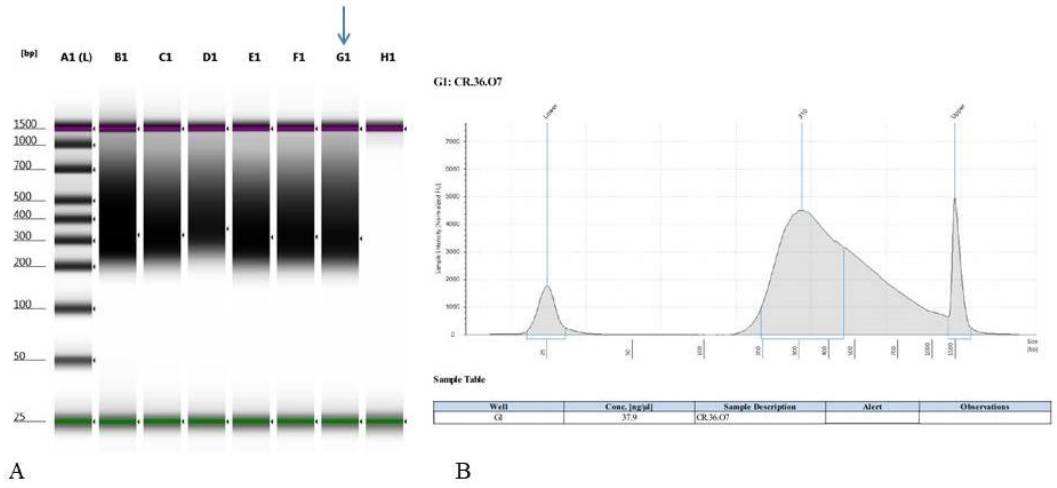


Figure 13. Image of D1000 Tape station. A) Gel image of the DNA analysis for samples from the NGS library workflow using the 2200 TapeStation system. B) Region views of the Agilent TapeStation Analysis Software of CR_36_O7.

For the hybridization step, 750 ng in 12 µl of each prepared library were added to 5 µl of SureSelect QXT Fast Blocker Mix and incubated into the thermal cycler with the following program: 35°C for 5 minutes, 65°C for 10 minutes, 65°C for 1 minute, 65°C for 1 minute and 37°C for 3 seconds; the last two steps were repeated for 60 cycles and 65°C hold. During the second step at 65°C, in the thermal cycler, to each sample were added 13 µl of mix containing: 0.5 µl of SureSelect RNase Block, 2 µl

of Capture library, 6 µl of SureSelect Fast Hybridization Buffer and 4.5 µl of Nuclease-free water. At this step, the samples were purified with streptavidin-coated magnetic beads so, for each hybridization sample 50 µl of Dynabeads MyOne Streptavidin (Thermo Fisher Scientific, Waltham, MA, USA), were washed three times with 200 µl of SureSelect Binding Buffer and, at the end of this steps, each hybridization sample was added to the 200 µl of washed streptavidin beads. The samples were incubated for 30 minutes at 1800 rpm at room temperature; then, the samples were kept in a magnetic separator. The supernatant was removed, and the beads were washed firstly with 200 µl of Sureselect Wash Buffer 1 and then three times with 200 µl of prewarmed Wash Buffer 2. At the last step, the beads were resuspended with 23 µl of Nuclease-free water. The SureSelect-enriched DNA libraries were PCR amplified using the appropriate pair of dual indexing primer; the PCR mix was formed by 10 µl of Herculare II 5x Reaction Buffer, 0.5 µl of 100 mM dNTP Mix, 1 µl of Herculase II Fusion DNA Polymerase and 13.5 µl of Nuclease-free water. Twenty-five µl of prepared PCR-mix were added to 1 µl of appropriate P7 dual index primer, to 1 µl of appropriate P5 dual index primer and to 23 µl of SureSelect-Enriched DNA libraries. The PCR was run into the thermal cycler following this program: 98°C for 2 minutes, 98°C for 30 seconds, 58°C for 30 seconds and 72°C for 1 minutes; the last 3 steps were repeated for a total of 14 cycles, 72°C for 5 minutes and 4°C hold. When the PCR amplification program was completed, strip tubes were briefly spin and the streptavidin-coated beads were removed placing the strip tubes on the magnetic stand at room temperature. The supernatant from each tube was recovered and each sample was purified with AMPure XP bead as described before.

To verify the quality of the enriched library and to quantify it, a High Sensitivity D1000 ScreenTape (Agilent Technologies, Santa Clara, CA, USA) was used, following manufacturer's instructions (Figure 14).

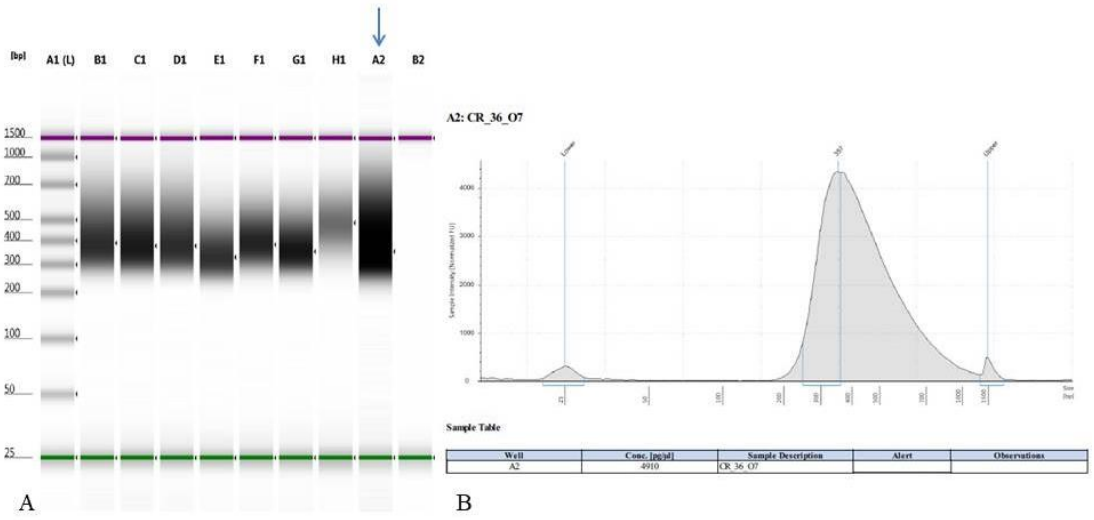


Figure 14. Image of High Sensitivity D1000 Tape station. A) Gel image of the DNA analysis for samples from the NGS library workflow using the 2200 TapeStation system. B) Region views of the Agilent TapeStation Analysis Software of CR_36_O7.

3.8 Next generation sequencing with Illumina instrument

Multiple samples were sequenced together, since they were univocally tagged by specific indexing primer during DNA library preparation. In particular, 11 paired-end (PE 150x2) sequencing runs, with 10 samples each, were carried out on a Miseq platform (Illumina Inc., San Diego, CA, USA) with MiSeq® Reagent Kit v2 (500 cycle). The number of samples to be sequenced per run was chosen based on the platform's maximum output and based on the calculation of the optimal theoretical coverage chosen by the researcher (150/200X for target regions).

Prior to pooling, DNA samples were quantified using a HS Qubit Fluorometer following the Invitrogen protocol. Based on their concentration and molarity, the

samples were diluted to 4 nM. Samples with different indexes were pooled to proceed with the library denaturation step following the Illumina protocol.

Eleven different runs were optimized for MiSeq System, 8 pM of the denatured final libraries pool was combined to 25% of 8 pM PhiX and were loaded into the MiSeq V2 reagent cartridge.

There are four important parameters to check during the sequencing run: the cluster density shows the number of clusters that have formed on the surface of the flowcell per square millimetre, clusters passing filter (PF%) shows the percentage of clusters that pass filters (this is a measure of the quality of libraries prepared), estimated yield (Mb) shows the projected number of bases called for the run, measured in megabases and the quality score (Q-score) that shows the average percentage of bases greater than Q30, which is a quality score measurement. A Q-score is a prediction of the probability of a wrong base call (Q30 = 1 in 1000).

3.9 Analysis with Alissa pipeline

At the end of sequencing runs a big quantity of information is produced and is important to understand the significance of all this data to make the right considerations. During the years, a lot of different software has been used but, all of them are based on common concept: the creation of a specific pipeline that can convert luminescence or fluorescence images acquired by the instruments into nucleotide sequences, called sequencing reads.

For this project we used Agilent's SureCall software (Agilent Technologies), a bioinformatic tool that align and call all the variants found. This software is made by two modules: Alissa Align & Call that align sequencing reads to Human genome

reference sequence calculate quality control metrics and identify variants. Then, variant call format (VCF) and QC files were uploaded in Alissa Interpret that annotate and analyse variants found in the first module. For each analysed subjects, all the identified variants were filtered in a stabilized pipeline, first based on the localization (exonic, intronic, UTR3', UTR5') then, they were filtered for the coding effect generated and, among them, we applied another filter to identify and manage variants based on the clinical significance (benign, likely benign, uncertain significance, likely pathogenic and pathogenic variants), see figure 15, using the ClinVar database (https://www.ncbi.nlm.nih.gov/clinvar/).

3.10 Library Preparation for the Whole Genome Sequencing with Nanopore- Oxford Strategy (ONT), third generation sequencing

Four of our samples (organoid, tumor and healthy paired tissues) were analysed for Whole genome sequencing with Nanopore Oxford protocol. For each sample a single library was prepared following Genomic DNA by ligation (Oxford Nanopore Technologies, Oxford, UK). In short, potential nicks in DNA and DNA ends were repaired in a combined step using a NEBNext FFPE DNA repair mix and NEBNext Ultra II End Repair/dA-Tailing Module (New England Biolabs) followed by AMPure bead purification and ligation of sequencing adaptors into prepared ends (Figure16A). More in the details, 1 µg of high molecular weight genomic DNA in 48 µl of Nuclease free-water was added to 3.5 µl of NEBNext FFPE DNA Repair Buffer, 2 µl of NEBNext FFPE DNA Repair Mix, 3.5 µl of Ultra II End-prep reaction buffer and 3 µl of Ultra II End-prep enzyme mix (New England Biolabs, Massachusetts,USA).

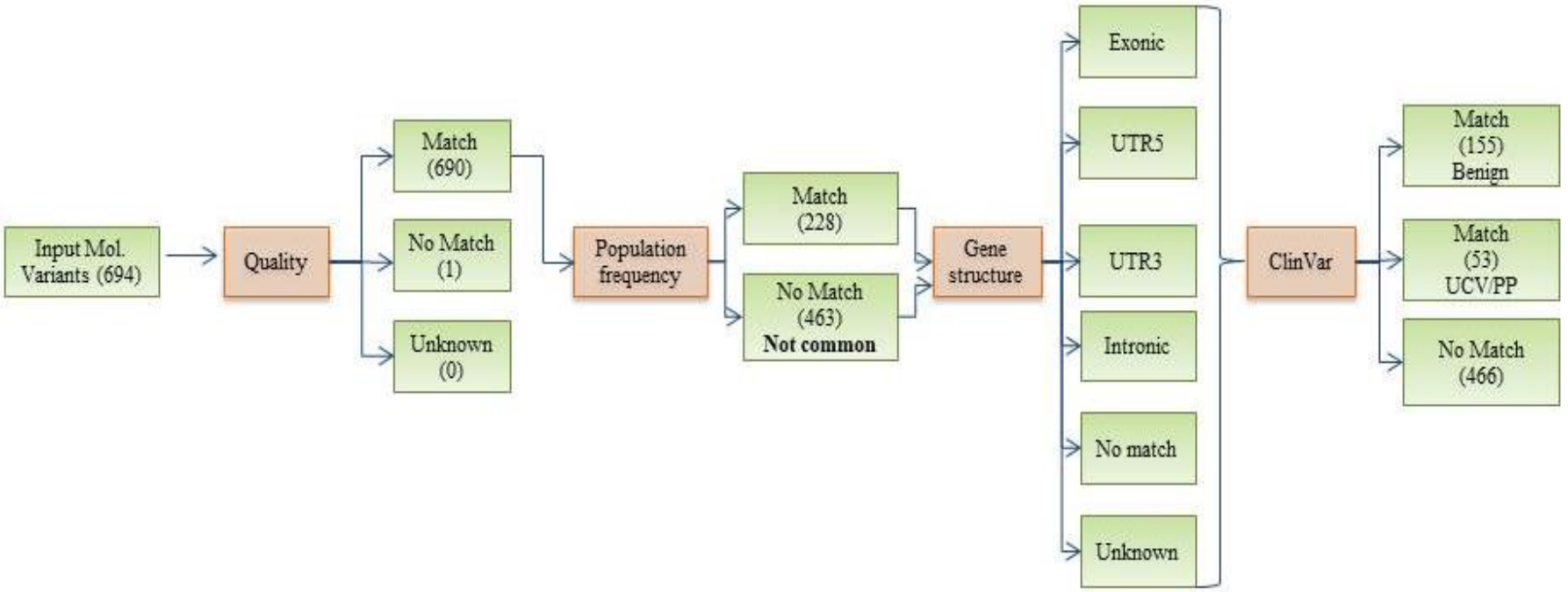


Figure 15. Alissa pipeline. The scheme by which we filtered the variants found with Illumina sequencing.

The sample was incubated into the thermal cycler at 20°C for 5 minutes and 65°C for 5 minutes. At the end of the incubation, at each sample were added 60 µl of AMPure XP beads, the obtained solution was incubated for 5 minutes at room temperature and the tubes were kept on the magnetic plate. The cleared supernatant was discarded; wash step was performed by adding 200 µl of 70% EtOH into each sample tube, the Ethanol was removed, and the step was repeated for a total of two washes. The tubes were air-dried with open lids at room temperature until the residual ethanol was evaporated; beads were resuspended with 61 µl of Nuclease-free water, incubated for 2 minutes at room temperature and in the final step the eluate was retained. The DNA derived by the previous step was used for the ligation of the adapter; 25 µl of Ligation Buffer, 10 µl of NEBNext Quick T4 DNA Ligase and 5 µl of Adapter Mix F were added to 60 µl of end-prepped DNA. The prepared solution was incubated for 10 minutes at room temperature, 40 µl of AMPure XP beads were added and another incubation step of 5 minutes was performed. At that point, the beads were washed twice with 250 µl of Long Fragment Buffer to retain only the fragment of 3kb or longer. The beads were eluted with 25 µl of Elution Buffer and after 10 minutes at 37°C the library was retained in a new 1.5 ml Eppendorf tube. Then, the libraries were sequenced with PtomethIon24 (Figure 16B).

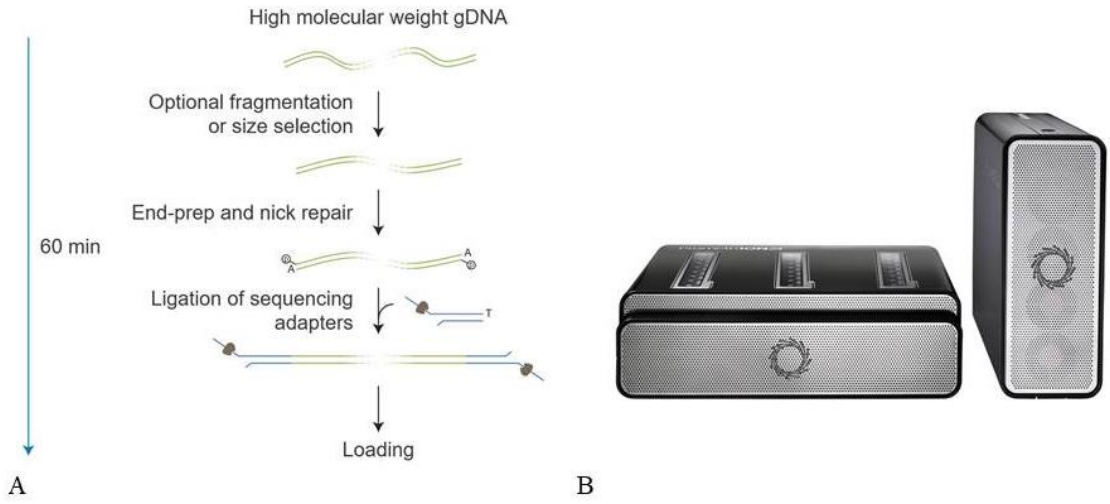


Figure 16. WGS with Oxford Nanopore Technology. A) Flow chart of the library preparation with ONT protocol (SQK-LSK110). B) An image of the PromethIon-24 sequencer.

3.11 Sequencing with PromethIon24- ONT Technologies

Single sample was sequenced with a PromethIon Flowcell (FLO-PRO002) following the manufacturer's instructions. In our case, we loaded on the flow cells the entire quantity of prepared library. The 24 µl of DNA library were added to 75 µl of Sequencing Buffer II and to 51 µl of Loading Beads II. The sample prepared as described before was uploaded through the inlet port of the flow cell. The run-time is 72 hours each. Base calling of the raw nanopore reads was performed using the Oxford Nanopore base caller Guppy on the PromethION computer device.

3.12 Immunofluorescence of Patient Derived Organoids (PDOs)

Immunofluorescence assay was performed in the stabilized PDTOs, and we followed the protocol described below. Seven days before, were plated 20 µl of matrigel contained organoids in a 24-multiwell. Media was changed at least every two days. When organoids are ready for fixation, media was removed from each well

and 600 µl of solution made by 300 µl of Acetone and 300 µl of MetOH were added to each well and incubated for 1 hour at room temperature. The solution was removed and 3 washing of 30 minutes each were performed with 1x FBS (Thermo Fisher Scientific, Waltham, MA, USA) and 0.1x Triton (Sigma-Aldrich, St. Louis, Missouri, US) in PBS. To block the reaction, the last washing was removed and, to each well were added 10x FBS, 0.1x Triton and PBS for 1 hour at room temperature. Rat monoclonal anti-Ki67 primary antibody diluted at 1:10, Phalloidin diluted at 1:10 and cytokeratin 20 diluted at 1:50 were added to each well and incubated overnight at 4°C. After overnight incubation, each well was washed 3 times for 30 min each with 300 µl of wash solution. Secondary antibody anti rabbit alexa fluor-549 and alexa fluor-488 (Thermo Fisher Scientific, Waltham, MA, USA) was prepared in wash Solution. To visualize DNA/nuclei, a solution of DAPI (1:1500) in 1xPBS was made and 300 µl were added to each well for 15 min. Then, wash step was performed with 1xPBS for 5 minutes at room temperature.

The slide was mounted on the specimen slide with 20 µl of Mowiol (Sigma-Aldrich, St. Louis, Missouri, US). Organoids were visualized by confocal microscope with LSM 980, Zeiss.

4. Results

4.1 Cultured and analyzed PDOs

All the patients enrolled (n=80) were analyzed for germline gene mutations; moreover, we established 11 PDOs from the enrolled patients (Table 3). For this subgroup (n=11) we performed the comparison of the four genomes: blood, PDOs, tumoral and paired heathy-tissue from the same patient (Figure 17).

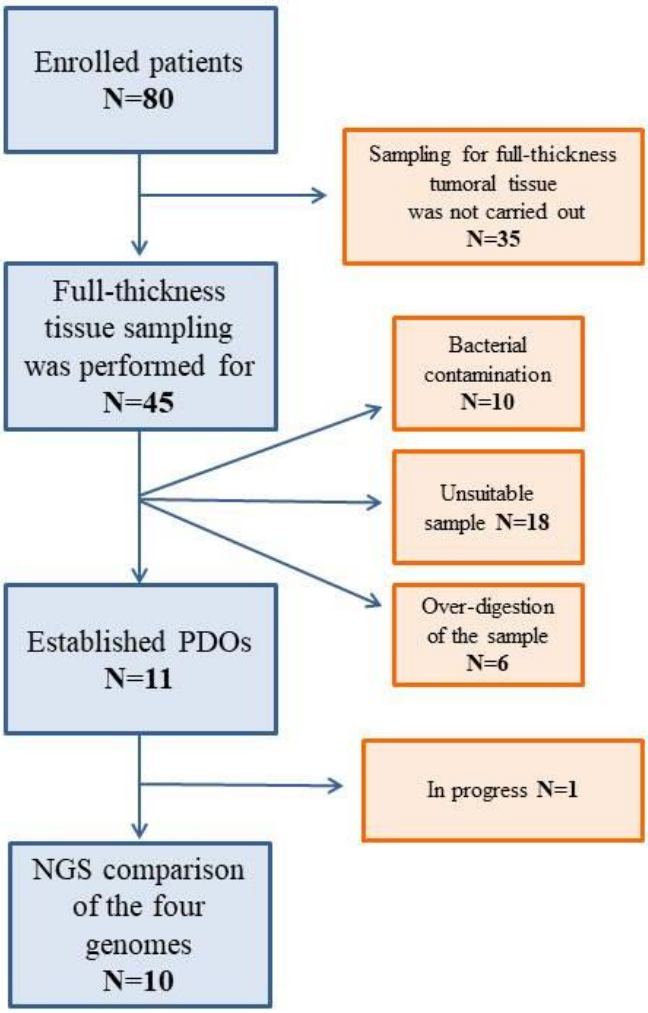


Figure 17. Diagram of the patients enrolled and analysed. The following figure shows the numbers of patients enrolled, and the established 11 PDOs.

Considering as above, also highlighted in figure 17, the stabilization process of the PDOs is tricky and depends on the starting tissue. If the tumor is not big enough, the full-thickness tissue is not carried out to not affect the diagnostic process of the patient; indeed, the sampling phase will affect the result. Furthermore, in our case, working with a colon tissue, the decontamination process is very important and must be carried out carefully. Indeed, all these reasons were elements of exclusion of part of the initially stabilized organoids.

Table 3. Coloncancer patients (n=11) from whom we established PDOs.

N	ID Patient	Male Female	Age	Other tumor	Other pathologies	Comparison of the 4 genomes
1	CR_32	F	83	NO	NO	YES
2	CR_34	M	58	Oesophageal adenocarcinoma	Type II diabetes mellitus, hypercholesterolemia, arterial hypertension	YES
3	CR_35	M	53	NO	-	YES
4	CR_36	F	73	NO	Hepatitis C, diabetes, arterial hypertension	YES
5	CO_03	F	64	Breast cancer	Type II diabetes mellitus, arterial hypertension	YES
6	CO_09	F	61	NO	Gastritis, duodenal ulcer	YES
7	CO_10	M	83	NO	Type II diabetes mellitus, chronic obstructive pulmonary disease, acute myocardial infarction	YES
8	CO_22	M	69	K lung, not small cell	Arterial hypertension	YES
9	CO_36	F	73	NO	Thyroid nodule	YES
10	CO_38	F	43	NO	Anemia	YES
11	CO_42	M	87	NO	Chronic obstructive pulmonary disease, acute myocardial infarction, chronic renal failure, rheumatoid arthritis, stroke, benign prostatic hyperplasia	In progress

4.2 Raw data of NGS

Eleven different runs were performed to analyze all the patients selected for the present study. In particular, all the 80 enrolled patients were analyzed for germline mutations and for these we performed 8 different runs with 10 samples each; moreover, we performed 3 runs with 10 samples each, for the comparison of PDOs, tumoral and paired heathy tissue from the same patient.

Different parameters can be checked during the sequencing run: the cluster density shows the number of clusters that have been formed on the surface of the flowcell per square millimeter, clusters passing filter (PF%) shows the percentage of clusters that pass filters (this is a measure of the quality of the prepared libraries), estimated yield (Mb) shows the number of bases called for the run, measured in megabases, and the quality score (Q-score) that shows the average percentage of bases greater than Q30, which is a quality score measurement. A Qscore is a prediction of the probability to get a wrong base call (Q30 = 1 in 1,000). For a final data related to the quality of sequencing runs the cluster density is an important factor in optimizing data quality and yield. The NGS procedure, as far as the relevant parameter, connected with sequencing runs were recognized as optimal measures which deserve high level of quality and sufficient presence of library pool.

For our 11 runs we obtained a Cluster density between 800-1000 K/mm2, Cluster passing filters between 91-97% and Q30 between 94.4-96.7.

The files produced are fastQ files and each of that are univocally assigned to each patient thanks to the index sequence, added during the library preparation, and can be exported from our platform BaseSpace Onsite (that provides the storage, and the sharing of NGS data) to be analyzed through Alissa.

All details and results concerning the variants are reported, as indicated below, in Table 4 and 5; for those, we showed the pathogenic variants in the somatic (PDOs, tumoral and paired-healthy tissue) and in the germinal (blood samples) line too. Those variants are reported in bioinformatics database such as Clinvar, and in *in-silico* tool that predict the pathogenicity on the basis of already published data.

4.3 Comparison of the mutational pattern of blood, PDOs, tumor and paired-healthy tissue in the same patient of multi-gene panel screening

After carrying out molecular analyses at the germinal level; as soon as the organoids were stabilized, from step 5 onwards, we performed the sequencing of the trio (PDOs, tumor and paired-healthy tissue) samples of the multi-gene panel with Illumina sequencing. In this phase, it was very important to understand the mutational pattern of organoids with respect to the tumoral tissue. Of course, with the multi-gene panel we found benign mutations (about 300-400 variants each) in common between the three samples of the same subject but, interestingly, the pathogenic mutations found in the PDOs were almost always confirmed at the level of the tumor tissue (Table 4). Moreover, the variants found, including the benign ones in almost always our patients were found in all the tissues but, sometimes, the variants can appear also at germinal level (in DNA extracted from blood); this means that all the variants should be considered somatic, and they appear during the time of the tumor growth.

In our PDOs samples were found 37 mutations: 23 pathogenic mutations reported in Clinvar database, 1 reported in Clinvar as Likely Pathogenic, 4 mutations reported as VUS in Clinvar database and predicted as Pathogenic after *in-silico* analyses and 9

mutations not reported in Clinvar but Pathogenic/Likely pathogenic after *in-silico* predictions. Of the 24 pathogenic/likely pathogenic mutations, 9 were missense and 15 nonsense; of the 4 VUS, 3 were missense and 1 nonsense; of the 8 variants not reported in Clinvar, 3 were missense variant and 6 nonsense variants.

Moreover, in three patients (CO_09, CO_36 and CO_38) we found three different variants, not reported in Clinvar database but predicted Pathogenic/Likely Pathogenic present both in the germinal line and at the somatic level.

As it can be noted from table 4 in several instances are present mutations in the organoid structures do not present in the tumor tissues; this is most likely due to the fact that in the organoids the cells containing mutations are more numerous respect to that present at the tumoral level; those, resulting at the sequencing level.

The cellular multiplicity in the organoids also determines a cloning of tumor cells which allow obtaining results that are more sufficiently valid for the robustness of the results obtained. This could therefore be a great advantage of this technology as, to date, mutational analyses at the somatic level may not highlight all the mutations present in tumors because, the proportion of tumor cells may be too small to be confirmed at the level of DNA sequencing. Instead, in the organoid, as there is cellular expansion and multiplication, this problem seems to be solved

Table 4. Variants found with multi-gene panel in four-genome comparison: blood (T0), PDOs (_O), tumoral (K1) and paired-healthy (S1) tissue from the same patients. For each patient ID the corresponding sample (T0, O, K1 and S1) was reported from which the DNA was extracted and subsequently sequenced. The number in the first line is consequential and it is indicated only when the variant is found for the first time, in the other samples the same variant is indicated with a hyphen (-).

N° of variants	ID of the patients	Sample	Gene	cDNA	Protein	Reference SNP ID number	ClinVar Cassitication	ACMG score	Varsome (Dann Score)
-	CR_32	T0				WT			
1		O7#	*PIK3CA*	c.3140A>G	p.His1047Arg	rs121913279	Pathogenic	Pathogenic	0.990
2			*TP53*	c.660T>G	p.Tyr220*	nr	NA	Pathogenic	0.989
-		K1	*TP53*	c.660T>G	p.Tyr220*	nr	NA	Pathogenic	0.989
-		S1				WT			
-	CR_34	T0				WT			
3		O5#	*APC*	c.4012C>T	p.Gln1338*	rs121913327	Pathogenic	UCV	0.967
-		K1	*APC*	c.4012C>T	p.Gln1338*	rs121913327	Pathogenic	UCV	0.967
-		S1				WT			
-	CR_35	T0				WT			
4		O7#	*MSH2*	c.2246A>G	p.Glu749Gly	nr	NA	Likely Pathogenic	0.999
5			*MSH6*	c.3700G>T	p.Glu1234*	nr	Pathogenic	Pathogenic	0.997
6			*PIK3CA*	c.137A>C	p.Lys46Thr	nr	NA	Likely Pathogenic	0.998
7			*PIK3CA*	c.353G>A	p.Gly118Asp	rs587777790	Likely Pathogenic	Pathogenic	0.997
8			*APC*	c.3340C>T	p.Arg1114*	rs121913331	Pathogenic	Pathogenic	0.997
9			*APC*	c.4508C>A	p.Ser1503*	nr	NA	Pathogenic	0.998
10			*APC*	c.6610C>T	p.Arg2204*	rs752654519	Pathogenic	Pathogenic	0.996
11			*APC*	c.6709C>T	p.Arg2237*	rs768922431	Pathogenic	Pathogenic	0.996
12			*APC*	c.7709C>A	p.Ser2570*	nr	NA	Pathogenic	0.995
13			*PMS2*	c.826G>T	p.Glu276*	nr	NA	Pathogenic	0.998

N° of variants	ID of the patients	Sample	Gene	cDNA	Protein	Reference SNP ID number	ClinVar Cassitication	ACMG score	Varsome (Dann Score)
14			*ATM*	c.748C>T	p.Arg250*	rs772821016	Pathogenic	Pathogenic	0.997
15			*BRCA2*	c.1528G>T	p.Glu510*	rs80358438	Pathogenic	Pathogenic	0.991
16			*BRCA2*	c.3922G>T	p.Glu1308*	rs80358638	Pathogenic	Pathogenic	0.994
17			*NF1*	c.7285C>T	p.Arg2450*	rs786202457	Pathogenic	Pathogenic	0.995
-			*MSH2*	c.2246A>G	p.Glu749Gly	nr	NA	Likely Pathogenic	0.999
-			*MSH6*	c.3700G>T	p.Glu1234*	nr	Pathogenic	Pathogenic	0.997
-			*PIK3CA*	c.137A>C	p.Lys46Thr	nr	NA	Likely Pathogenic	0.998
-			*PIK3CA*	c.353G>A	p.Gly118Asp	rs587777790	Likely Pathogenic	Pathogenic	0.997
-			*APC*	c.3340C>T	p.Arg1114*	rs121913331	Pathogenic	Pathogenic	0.997
-		K1	*APC*	c.4508C>A	p.Ser1503*	nr	NA	Pathogenic	0.998
-			*APC*	c.6610C>T	p.Arg2204*	rs752654519	Pathogenic	Pathogenic	0.996
-			*APC*	c.6709C>T	p.Arg2237*	rs768922431	Pathogenic	Pathogenic	0.996
-			*APC*	c.7709C>A	p.Ser2570*	nr	NA	Pathogenic	0.995
-			*PMS2*	c.826G>T	p.Glu276*	nr	NA	Pathogenic	0.998
-			*ATM*	c.748C>T	p.Arg250*	rs772821016	Pathogenic	Pathogenic	0.997
-			*BRCA2*	c.1528G>T	p.Glu510*	rs80358438	Pathogenic	Pathogenic	0.991
-			*BRCA2*	c.3922G>T	p.Glu1308*	rs80358638	Pathogenic	Pathogenic	0.994
-			*NF1*	c.7285C>T	p.Arg2450*	rs786202457	Pathogenic	Pathogenic	0.995
-		S1				WT			
-		T0				WT			
18		O7[#]	*APC*	c.2626C>T	p.Arg876*	rs121913333	Pathogenic	Pathogenic	0.997
19	CR_36		*TP53*	c.637C>T	p.Arg213*	rs397516436	Pathogenic	Pathogenic	0.998
-		K1	*APC*	c.2626C>T	p.Arg876*	rs121913333	Pathogenic	Pathogenic	0.997
-			*TP53*	c.637C>T	p.Arg213*	rs397516436	Pathogenic	Pathogenic	0.998
-		S1				WT			

N° of variants	ID of the patients	Sample	Gene	cDNA	Protein	Reference SNP ID number	ClinVar Cassitication	ACMG score	Varsome (Dann Score)
-	CO_03	T0				WT			
20		O6#	*MLH1*	c.37G>A	p.Glu13Lys		VUS	VUS/Likely Pathogenic	0.999
21			*APC*	c.697C>T	p.Gln233*	rs1554074772	Pathogenic	Pathogenic	0.998
22			*APC*	c.4585C>T	p.Gln1529*	rs1554085992	Pathogenic	Pathogenic	0.997
23			*TP53*	c.733G>A	p.Gly245Ser	rs28934575	Pathogenic	Pathogenic	0.998
-		K1				WT			
-		S1				WT			
24	CO_09	T0	*ATM*	c.6067G>A	p.Gly2023Arg	rs11212587	VUS	Likely Pathogenic	0.999
25		O5#	*APC*	c.4288delA	p.Thr1430Profs*43	nr	NA	Pathogenic	-
26			*TP53*	c.524G>A	p.Arg175His	rs28934578	Pathogenic	Pathogenic	0.999
-			*ATM*	c.6067G>A	p.Gly2023Arg	rs11212587	VUS	Likely Pathogenic	0.999
-		K1	*TP53*	c.524G>A	p.Arg175His	rs28934578	Pathogenic	Pathogenic	0.999
-		S1				WT			
-	CO_10	T0				WT			
27		O5#	*APC*	c.4348C>T	p.Arg1450*	rs121913332	Pathogenic	Pathogenic	0.997
28			*TP53*	c.733G>A	p.Gly245Ser	rs28934575	Pathogenic	Pathogenic	0.998
-		K1	*APC*	c.4348C>T	p.Arg1450*	rs121913332	Pathogenic	Pathogenic	0.997
-			*TP53*	c.733G>A	p.Gly245Ser	rs28934575	Pathogenic	Pathogenic	0.998
-		S1				WT			

N° of variants	ID of the patients	Sample	Gene	cDNA	Protein	Reference SNP ID number	ClinVar Cassitication	ACMG score	Varsome (Dann Score)
-	CO_22	T0				WT			
29		O7[#]	*TP53*	c.584T>C	p.Ile195Thr	rs760043106	Pathogenic	Pathogenic	0.992
-		K1	*TP53*	c.584T>C	p.Ile195Thr	rs760043107	Pathogenic	Pathogenic	0.993
-		S1				WT			
30	CO_36	T0	*BMPR1A*	c.1245A>C	p.Glu415Asp	rs786204235	VUS	Likely Pathogenic	0.986
31		O7[#]	*PIK3CA*	c.1624G>A	p.Glu542Lys	rs121913273	Pathogenic	Pathogenic	0.999
32			*APC*	c.637C>T	p.Arg213*	rs587781392	Pathogenic	Pathogenic	0.997
33			*APC*	c.4132C>T	p.Gln1378*	rs121913329	Pathogenic	Pathogenic	0.997
34			*TP53*	c.817C>T	p.Arg273Cys	rs121913343	Pathogenic	Pathogenic	0.998
-			*BMPR1A*	c.1245A>C	p.Glu415Asp	rs786204235	VUS	Likely Pathogenic	0.986
-		K1	*BMPR1A*	c.1245A>C	p.Glu415Asp	rs786204235	VUS	Likely Pathogenic	0.986
-		S1	*BMPR1A*	c.1245A>C	p.Glu415Asp	rs786204235	VUS	Likely Pathogenic	0.986
35	CO_38	T0	*SMARCA4*	c.3854T>A	p.Leu1285*	nr	NA	Pathogenic	0.987
36		O7[#]	*PIK3CA*	c.2176G>A	p.Glu726Lys	rs867262025	Pathogenic	Pathogenic	0.997
37			*DICER1*	c.5176C>T	p.Gln1726*	nr	VUS	Pathogenic	0.998
-			*SMARCA4*	c.3854T>A	p.Leu1285*	nr	NA	Pathogenic	0.987
-		K1	*SMARCA4*	c.3854T>A	p.Leu1285*	nr	NA	Pathogenic	0.987
-		S1	*SMARCA4*	c.3854T>A	p.Leu1285*	nr	NA	Pathogenic	0.987

[#] **O5,O6, O7**: DNA somatic variant at the fifth, sixth and seventh passage of organoid splitting; **n.r.:** not reported; **NA**: Not available.

4.4 Variants found in the germline DNA of enrolled patients

In the blood of our cohort of patients (n=80), we found 10 patients (12.5%) with Pathogenic or Likely Pathogenic mutations. Totally, we found 14 variants with clinical significance in genes related to CRC predisposition.

Moreover, another information appears to be relevant for the study approached in this thesis; in almost all case where pathogenic variants have been found, there are change of amminoacid, thus, indicating missense mutation (n=8).

Indeed, in our patients 8 variants were missense, 5 were nonsense and 1 was frameshift, a single base duplication, that produce a shift of the reading frame resulting in a premature termination codon (see table 5). Moreover, four of these patients show double mutations; indeed, among them, three have familiarity with oncological diseases and for one (CO_42) that information was missed. Two patients show 1 Pathogenic/Likely Pathogenic and 1 UCV variant; in one (CO_30), instead, there were 2 Pathogenic/Likely Pathogenic variants reported in Clinvar database; in the last patient (CO_42) there were 2 mutations, 1 in *PPM1D* gene not Available on Clinvar database, but predicted to be pathogenic in Varsome and 1 UCV variant. For this sample (CO_42) we are waiting for the results of the comparison of the genomes in PDOs, tumoral and paired-healthy tissue.

The clinical significance of all the variants (n= 14) was assessed based on Clinvar database, and, furthermore, for all the variants found we also carried out *in silico* prediction on bioinformatics tools (Varsome); we considered the values assigned by the ACMG criteria and the DANN score reported in this tool.

Only for one patient, we found a mutation not reported in Clinvar database, but predicted to be Pathogenic for ACMG criteria; this variant in *SMARCA4* gene is a nonsense variant.

4.5 Raw data for WGS with Oxford Nanopore Technology

For four of our PDOs, we have already carried out the WGS with Nanopore technology; for the others, the experiments are currently in progress.

Twelve different runs were performed with PromethIon24 and, different parameters were checked during the entire time of the experiment (usually 72h).

The flow cell check must be carried out before starting the run, as more than 5000 pores must be available in each flow cell to have a good yield. At the end of the run different parameters can be checked: the estimated bases indicate the amount of bases sequenced in the sample; obviously, it is important to distinguish between passed and failed bases and, those failed must be between 10-15% of the total. The evaluation of the estimated N50 must always be done, the *N50* is defined as the sequence length of the shortest contig at 50% of the total genome length. In our case, the goal is to obtain long reads (about 10,000 bp), which allow us to carry out specific bioinformatic analyses, not always possible with the reads obtained with the Illumina strategy (about 400 bp). All these parameters have been evaluated in our runs, as shown in the table 6.

Currently, methylome analyses, genome-wide variant calling and rearrangement analyses are ongoing. These will be necessary to better compare the genomes deriving from PDOs, tumor and paired-healthy tissue.

Table 5. The 14 Pathogenic or Likely Pathogenic variants found in the patients enrolled and analyzed for a multi-gene panel in germinal line.

N° of variants	Patient ID	Gene	cDNA	Protein	Reference SNP ID number	ClinVar Cassitication	ACMG score	Varsome (Dann Score)
1	CR_11	*MUTYH*	c.1187G>A	p.Gly396Asp	rs36053993	Pathogenic	Pathogenic	0.998
2	CR_30	*CHEK2*	c.1046G>C	p.Gly349Ala	rs587780192	Likely pathogenic	Likely pathogenic	0.998
3	CR_33	*APC*	c.3920T>A	p.Ile1307Lys	rs1801155	Likely pathogenic	Benign	0.89
4	CO_15	*MSH6*	c.3261dupC	p.Phe1088Leufs*5	rs1114167705	Pathogenic	Pathogenic	-
5	CO_17	*RNASEL*	c.793G>T	p.Glu265*	rs74315364	Likely pathogenic	Pathogenic	0.994
6		*MSH6*	c.3260C>A	p.Pro1087His	rs63750753	Uncertain significance	Uncertain significance	0.996
7	CO_27	*MUTYH*	c.536A>G	p.Tyr179Cys	rs34612342	Pathogenic	Pathogenic	0.997
8		*APC*	c.4073C>T	p.Ala1358Val	rs730881249	Uncertain significance	Uncertain significance	0.999
9	CO_29	*CHEK2*	c.1556C>T	p.Thr519Met	rs142763740	Likely pathogenic	Pathogenic	0.999
10	CO_30	*RNASEL*	c.793G>T	p.Glu265*	rs74315364	Likely pathogenic	Pathogenic	0.994
11		*ATM*	c.8977C>T	p.Arg2993*	rs770641163	Pathogenic	Pathogenic	0.989
12	CO_38	*SMARCA4*	c.3854T>A	p.Leu1285*	n.r.	NA	Pathogenic	0.987
13	CO_42	*PPM1D*	c.1403C>G	p.Ser468*	rs375975790	NA	Pathogenic	0.993
14		*POLE*	c.2263G>A	p.Val755Met	rs1222472060	Uncertain significance	Uncertain significance	0.999

n.r.: not reported; NA: Not available

Table 6. Raw data of the WGS with Oxford Nanopore Technology.

Raw data of WGS with Oxford Nanopore Technology					
N°	ID of the patient	Sample	Reads generated	Estimated Bases	Estimated N50
1		O6[#]	29.42 M	184.05 Gb	9.13 kb
2	CO_03	K1	17.54 M	104.46 Gb	7.68 kb
3		S1	26.72 M	143.71 Gb	7.69 kb
4		O5[#]	22.46 M	167.32 Gb	9.83 kb
5	CO_09	K2	26.96 M	141.56 Gb	6.69 kb
6		S1	39.03 M	144.86 Gb	4.82 kb
7		O5[#]	27.07 M	151.53 Gb	8.42 kb
8	CO_10	K1	26.47 M	148.73 Gb	7.93 kb
9		S1	22.94 M	121.51 Gb	6.98 kb
10		O6[#]	6.69 M	94.78 Gb	29.13 Kb
11	CO_22	K1	26.4 M	108.05 Gb	5.26 kb
12		S1	30.84 M	137.28 Gb	5.62 kb

[#] **O5, O6,**: DNA extracted from the fifth and sixth passage of organoid splitting.

4.6 Stabilization of PDOs from patient affected by CRC

With the protocol optimized in our laboratory we established 11 PDOs from patients affected by colorectal cancer (Figure 18); then, we were able to culture our PDOs up to numerous passages (p20) and no structural changes were seen after carefully observation under the microscope (e.g., for one of the PDOs see figure 19). This could mean that the organoids are quite stable and in turn it indicates the possibility of utilizing this system for successive analysis, including drug screening.

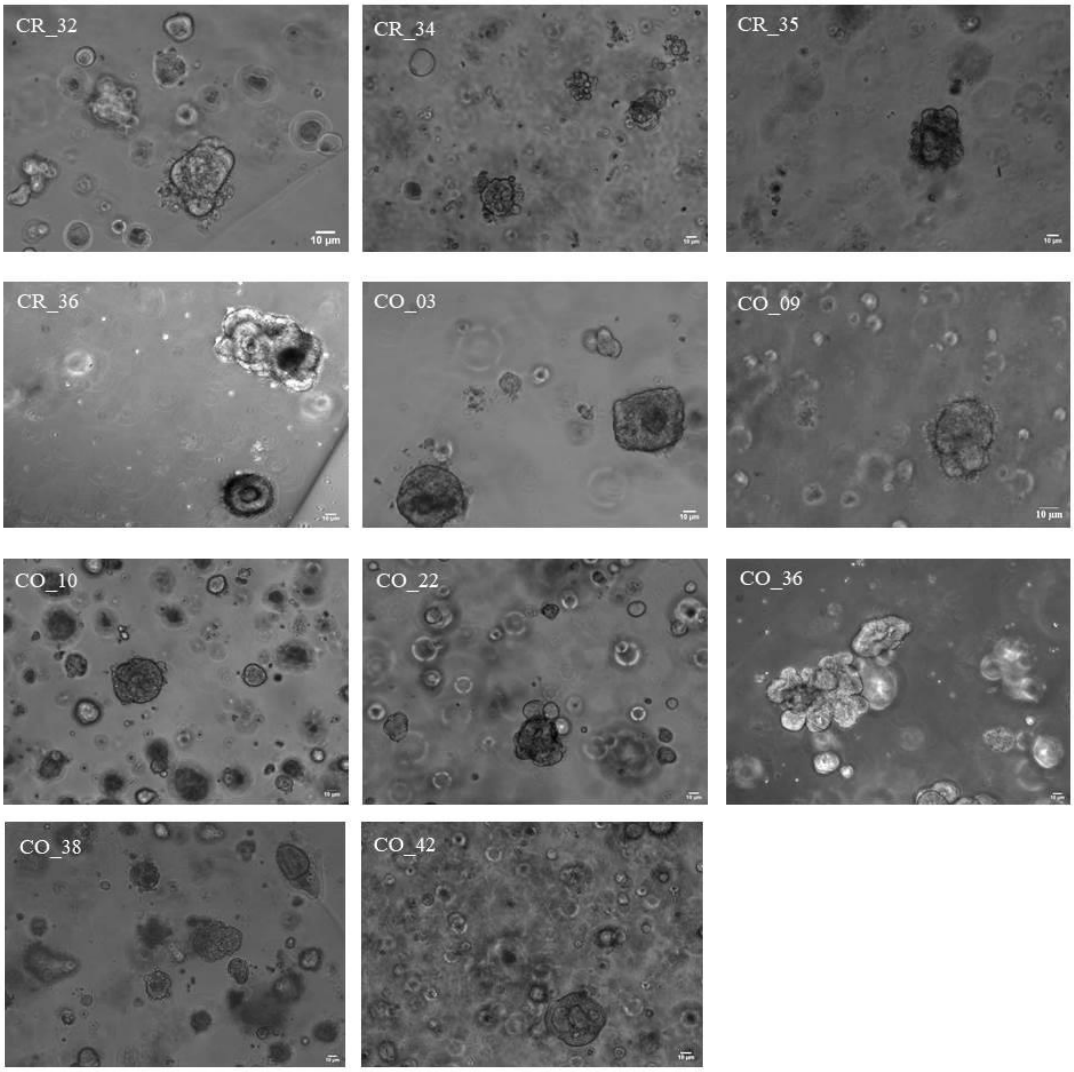


Figure 18. Images of the eleven established PDOs (scale bar: 10 μm). For each PDOs we acquired brightfield images with the LEICA-DMI4000 microscope, before carrying out the next split. The images shown of all PDOs are in the 6th passage.

The PDOs were always observed under the microscope before each split. During the various phases of growth, we follow them first, with DMI4000-Leica inverted optical microscope to acquire images of our organoids in brightfield (Figure 19).

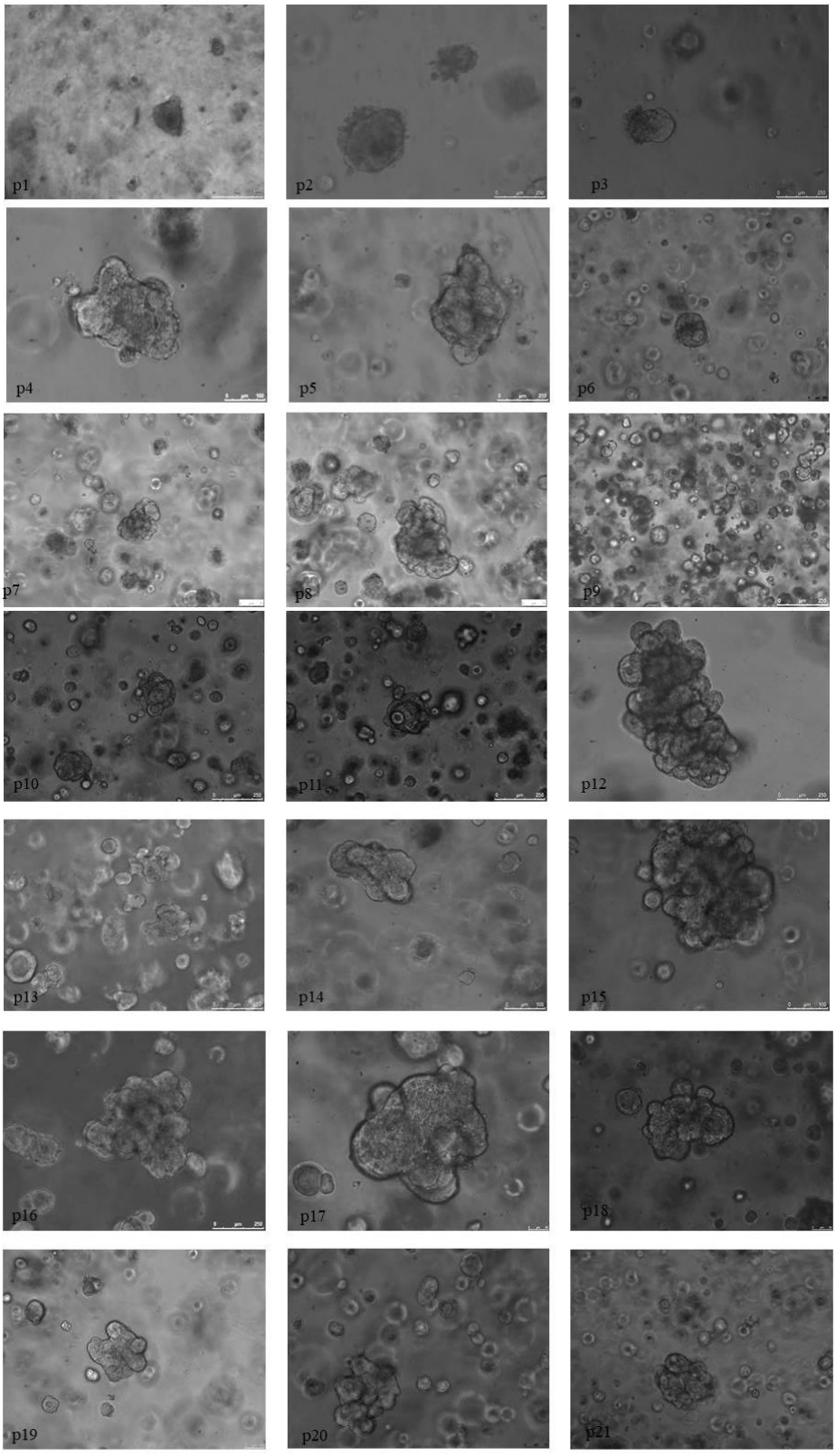


Figure 19. Twenty-one passages of one of established PDOs. The images show all the passages from p1 to p21 of the same PDOs (CO_22), scale bar 250µm.

4.7 Advanced light microscopy of our PDOs

To better characterize our PDOs, we carried out advanced microscopy analyses using the Cell Discoverer 7 - Zeiss which is an automated system that allow us to acquire high-resolution brightfield images (Figure 20). In was a very useful tool which highlighted the three-dimensionality of the organoids, also allowing us to clearly recognize and reconstruct the aggregative and cystic organoids as showed in figure 20.

Through this system we were able to distinguish two different structures present, well known also in literature: the cystic organoids and the aggregative ones.

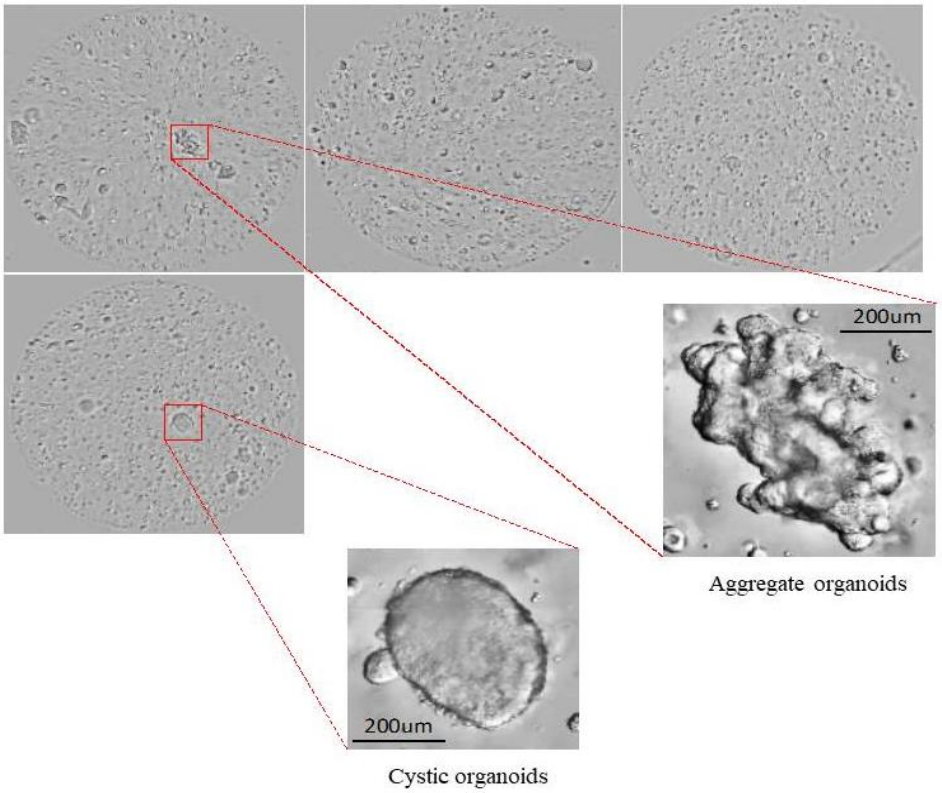


Figure 20. Reconstruction of organoids (CO_22), by advanced microscopy analysis. The gray squares show the matrigel domes in which the organoids are contained, then, through suitable software, the images are processed to reconstruct the structure of the single organoids.

Moreover, we performed immunofluorescence assay on our PDOs to verify the structures present in our organoids. Thanks to the use of specific antibodies such as Ki67, phalloidin and cytokeratin 20 (Figure 21) we have understood the internal structure of organoids. Indeed, we chose 3 different antibodies that localized in different purchased cells: Ki-67 localized at nuclear level; CK-20 and phalloidin both localized at cytoskeleton level.

Obviously, the presence of the Ki-67 shows a correlation with the clinical course of the neoplastic disease, on the other hand, the positivity to CK-20 confirms the presence of typical cell of colon adenocarcinomas.

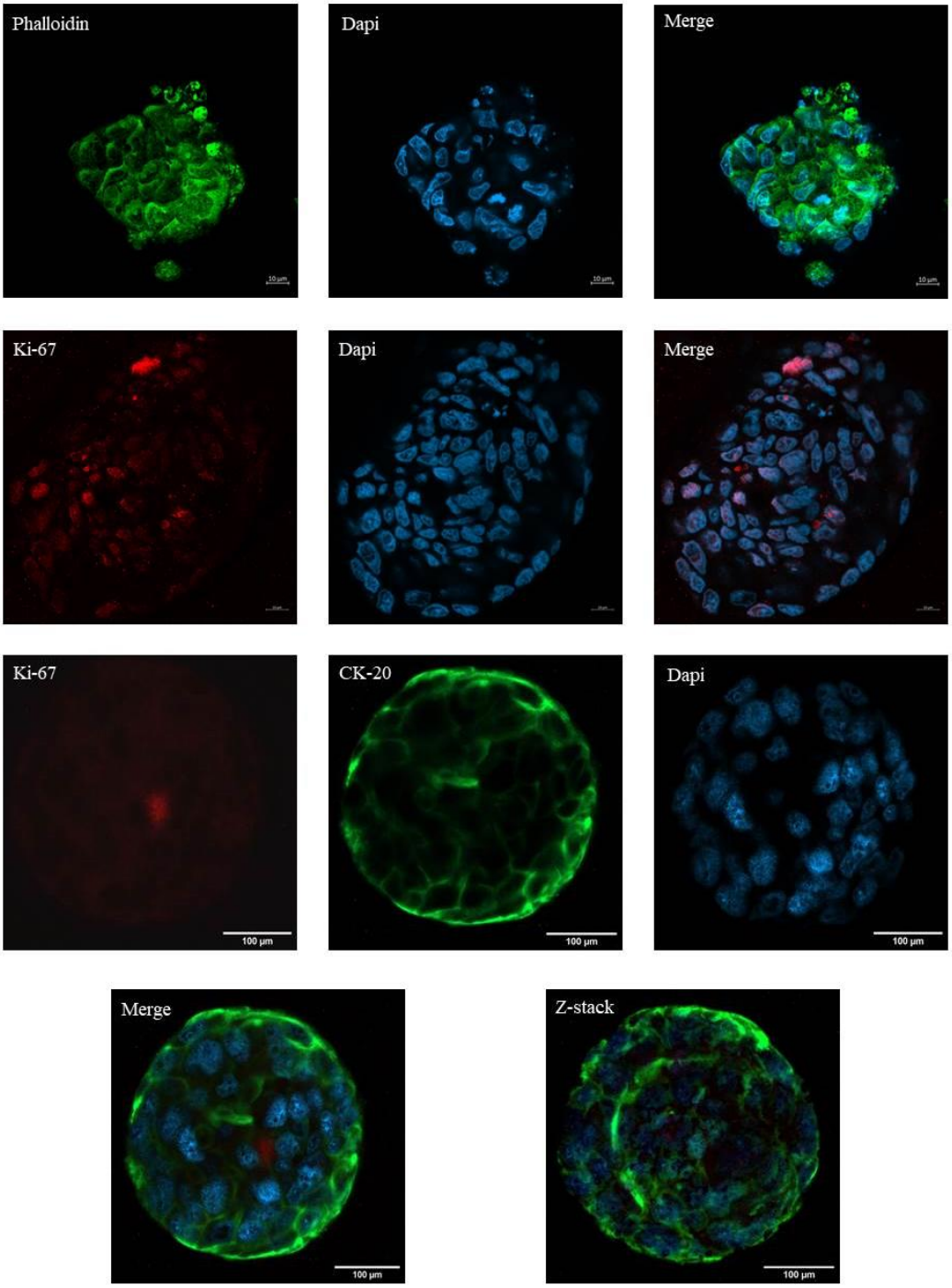


Figure 21. Immunofluorescence assay in PDO. CO_22. The PDOs was marked with Phalloidin (green), Ki-67 (red), CK-20 (green) and Dapi (blue). Moreover, we performed Z-stack (lower figure) to better reconstruct the PDO.

5. Discussion

Colorectal cancer is one of the most prevalent and heterogenous cancer worldwide. Patients differ in their response to treatment due to alterations at germinal and somatic levels [57]. Currently, various studies are underway to identify new mutations potentially involved in the etiopathogenesis of the tumor and in the response to chemotherapy [58]. The identification of new germline pathogenic mutations in genes associated with CRC may reveal at-risk subjects that are candidate for targeted therapy [58, 59].

Patient-derived organoids can be used to evaluate the pre-treatment effect with post-treatment, and it is also powerful tool that links patient-specific genetic and phenotypic information to preclinical and clinical drug response. The PDO model is becoming increasingly used to study tumor etiopathogenesis. One of the reasons is to evaluate the possibility of using a “patient-matched” drug approach in which patients receive the most appropriate drug for their condition. In fact, genomic data combined with drug screening data could lead to better cancer care.

Organoids are an important area of research because they recapitulate many functional and genetic features of the tissues from which they derived. Among the advantages of PDOs is that they remain stable throughout the freeze-thawing passages, thereby enabling high-throughput drug screening. Moreover, PDOs can be used to assess genomic landscapes, gene expression and transcriptome profiles of their parental tumors, even after long-term culture [59, 60].

The aim of this thesis is created PDOs from patient affected by CRC, and then compare the molecular pattern of PDOs to the germinal and the somatic line of

patient's DNA. The feasibility of developing CRC organoids is an important goal of this thesis, and it was demonstrated also by molecular analyses. The finding of pathogenic mutations in the PDOs and in the tumor tissues is a validation of the method we used. Our results and those of others confirm genotypic heterogeneity associated with predisposition to cancers commonly identified by germline genetic testing.

Twelve percent of our patients (n= 80) had germline gene mutations associated with CRC; 40% of them carried pathogenic mutations in the genes responsible for the familial forms. Two patients had mutations in the *MUTYH* responsible for the hereditary of MAP; 1 patient had a mutation in the *MSH6*, related to LS and finally, 1 patient had a mutation in *APC* related to FAP. The other 6 patients had pathogenic mutations in genes associated with breast and ovarian cancer, prostate and other syndromes related to CRC such as *CHEK2, ATM, POLE, RNASEL* and *SMARCA4* genes (see Table 3). Interestingly, we identified a mutation in one of our patients (ID: CO_38) that is not reported in the Clinvar database but predicted to be pathogenic according to *in-silico* tools. This mutation is in *SMARCA4,* and it creates a premature stop in the translation of mRNA into protein (p. Leu1285*). *SMARCA4* plays a role in the oncogenesis of various tumors. At somatic level missense point mutations in *SMARCA4* mapping to the ATPase domain thereby contributing to oncogenesis and/or epigenetic plasticity [61]. The finding of *SMARCA4* variant at both somatic and germinal level confirms the pathogenicity predicted by the bioinformatics tool.

An important aspect yet to be evaluated is to verify whether the differences we identified between the somatic and geminal lines can be applied in the diagnosis and

prognosis of patients [62]. The validation of possible biomarkers in DNA and/or in some species of RNA, for instance long noncoding RNA, short RNA, miRNA and circRNA can be used to stratify patients. One of the first biomarkers found was for breast cancer and led to notable advances in the treatment of patients. Unfortunately, many types of cancer do not have or have very few biomarkers, and this means a great challenge for future research.

Predicting biomarkers can be useful also at clinical level because their identification could help in the decision-making and management of patients. Notably, *KRAS* mutations are associated with increased risk of metastatic CRC and with a worse survival versus wild-type subjects [63]. In fact, *KRAS* mutant phenotypes respond poorly to anti-EGFR therapy [64]. Furthermore, *BRAF* mutations (i.e., V600E) are associated with reduced survival and up to 50% of patients have a lower survival versus wild-type subjects [63, 65-66]. Besides predicting disease progression, biomarkers can be used to tailor treatment according to molecular subtype. A second category of biomarkers are related to alterations in the germline genes. For example, the *VEGFA* polymorphism (rs833061) is associated with better survival in CRC patients treated with FOLFIRI and bevacizumab than in wild-type subjects.

In conclusion, the study of new biomarkers will lead to advances in the field of predictive medicine and in the management of patients. Indeed, using a personalized model like the original tumor can accurately predict the response of patients to treatment. Lastly, PDOs have advantages, challenges, and potential in individualized treatment.

6. Conclusion and future prospective

We established 10 CRC 3D in vitro organoid models that grow robustly for more than 6 months. Our four-step strategy revealed the presence of the pathogenetic variants in organoids and tumor tissues. It also confirms that the cells in *in-vitro* culture resemble the morphology of the tumor from which they derive, also at different passages. Obviously, the advantages of PDOs are their potential in cancer treatments and improvement of patient outcomes. Currently, we are screening drug related to the mutation found on our PDOs.

The aim of our study was to construct and characterize organoids derived from tumor taken freshly from tissues of patients. We are now looking for links between PDOs and the neoplastic tissue of each patient. Moreover, this approach results in: (i) a greater amount of material on which to perform pathophysiological analyses versus the tissue biopsied in clinical practice, and (ii) to study in greater detail therapeutic aspects, and to identify the most effective pharmacological approach to each patient.

We now plan to clarify the term “precision medicine”. This is devoted to beneficial aspects toward the disease, and the term usually refers to single alterations at molecular level (see DNA variants in tumors and/or genetic disease). “Precision medicine” in several instances may equate to “personalized medicine” since it is directed to a specific patient. However, “personalized medicine” is not “precision medicine” because it concerns the cure of each single patient often in its multi-morbidity aspects. In this context, the formation of a 3D culture of single neoplastic tissue to cultivate and characterize them is also of the utmost relevance.

Notably, we considered CRC to be a model neoplasia. Consequently, we characterized morphologically the PDOs and we sequenced the 58 genes associated with CRC in a single panel by NGS technology. We have also started a WGS approach with a third-generation sequencing technology (Oxford Nanopore Technology) that generates the whole sequence of the DNA obtained with our four-sample comparative analyses within the same subject. The ONT technology produces furthermore a complete epigenetic (methylome) analysis. Consequently, it is possible to analyze the CpG DNA regions and identify and compare which are more or less methylated.

In conclusion, this study opens new perspectives in the field of cancer research and in the treatment of patients affected by CRC.

7. References

1. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011 Mar 4; 144:646-74.

2. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin. 2021 May; 71:209-249.

3. Yamagishi H, Kuroda H, Imai Y, Hiraishi H. Molecular pathogenesis of sporadic colorectal cancers. Chin J Cancer. 2016 Jan 6.

4. Bailey CE, Hu CY, You YN, Bednarski BK, Rodriguez-Bigas MA, Skibber JM, Cantor SB, Chang GJ. Increasing disparities in the age-related incidences of colon and rectal cancers in the United States, 1975-2010. JAMA Surg. 2015 Jan;150(1):17-22.

5. Siegel RL, Fedewa SA, Anderson WF, Miller KD, Ma J, Rosenberg PS, Jemal A. Colorectal Cancer Incidence Patterns in the United States, 1974-2013. J Natl Cancer Inst. 2017 Aug 1.

6. Valle L, Vilar E, Tavtigian SV, Stoffel EM. Genetic predisposition to colorectal cancer: syndromes, genes, classification of genetic variants and implications for precision medicine. J Pathol. 2019 Apr; 247:574-588.

7. Yu H, Hemminki K. Genetic epidemiology of colorectal cancer and associated cancers. Mutagenesis. 2020 Jul 11; 35:207-219.

8. Siegel RL, Miller KD, Goding Sauer A, Fedewa SA, Butterly LF, Anderson JC, Cercek A, Smith RA, Jemal A. Colorectal cancer statistics, 2020. CA Cancer J Clin. 2020 May; 70:145-164.

9. Howlader N, Noone AM, Krapcho M, Miller D, Brest A, Yu M, Ruhl J, Tatalovich Z, Mariotto A, Lewis DR, Chen HS, Feuer EJ, Cronin KA (eds). SEER Cancer Statistics Review, 1975-2018, National Cancer Institute.

10. Cheng LK, O'Grady G, Du P, Egbuji JU, Windsor JA, Pullan AJ. Gastrointestinal system. Wiley Interdiscip Rev Syst Biol Med. 2010 Jan-Feb; 2:65-79.

11. Gerard J. Tortora, Bryan H. Derrickson. Principles of Anatomy and Physiology. 12th Edition.

12. Reed KK, Wickham R. Review of the gastrointestinal tract: from macro to micro. Semin Oncol Nurs. 2009 Feb; 25:3-14.

13. Greenwood-Van Meerveld B, Johnson AC, Grundy D. Gastrointestinal Physiology and Function. Handb Exp Pharmacol. 2017; 239:1-16.

14. Thibodeau GA, Patton KT. Anatomy and physiology. St.Louis, MO: Mosby; 2002. pp. 1354-1372.

15. American Cancer Society. Colorectal Cancer Facts & Figures 2020-2022. Atlanta: American Cancer Society; 2020.

16. Brenner H, Kloor M, Pox CP. Colorectal cancer. Lancet. 2014 Apr 26; 383: 1490-1502.

17. Mauri G, Sartore-Bianchi A, Russo AG, Marsoni S, Bardelli A, Siena S. Early-onset colorectal cancer in young individuals. Mol Oncol. 2019 Feb; 13:109-131.

18. Johnson CM, Wei C, Ensor JE, Smolenski DJ, Amos CI, Levin B, Berry DA. Meta-analyses of colorectal cancer risk factors. Cancer Causes Control. 2013 Jun; 24:1207-22.

19. Wang X, O'Connell K, Jeon J, Song M, Hunter D, Hoffmeister M, Lin Y, Berndt S, Brenner H, Chan AT, Chang-Claude J, Gong J, Gunter MJ, Harrison TA, Hayes RB, Joshi A, Newcomb P, Schoen R, Slattery ML, Vargas A, Potter JD, Le Marchand L, Giovannucci E, White E, Hsu L, Peters U, Du M. Combined effect of modifiable and non-modifiable risk factors for colorectal cancer risk in a pooled analysis of 11 population-based studies. BMJ Open Gastroenterol. 2019 Dec 2.

20. Thrift AP, Gong J, Peters U, Chang-Claude J, Rudolph A, Slattery ML, Chan AT, Esko T, Wood AR, Yang J, Vedantam S, Gustafsson S, Pers TH; GIANT Consortium, Baron JA, Bezieau S, Küry S, Ogino S, Berndt SI, Casey G, Haile RW, Du M, Harrison TA, Thornquist M, Duggan DJ, Le Marchand L, Lemire M, Lindor NM, Seminara D, Song M, Thibodeau SN, Cotterchio M, Win AK, Jenkins MA, Hopper JL, Ulrich CM, Potter JD, Newcomb PA, Schoen RE, Hoffmeister M, Brenner H, White E, Hsu L, Campbell PT. Mendelian randomization study of height and risk of colorectal cancer. Int J Epidemiol. 2015 Apr; 44:662-72.

21. Botteri E, Iodice S, Bagnardi V, Raimondi S, Lowenfels AB, Maisonneuve P. Smoking and colorectal cancer: a meta-analysis. JAMA. 2008 Dec 17; 300:2765-78.

22. Fedirko V, Tramacere I, Bagnardi V, Rota M, Scotti L, Islami F, Negri E, Straif K, Romieu I, La Vecchia C, Boffetta P, Jenab M. Alcohol drinking and colorectal cancer risk: an overall and dose-response meta-analysis of published studies. Ann Oncol. 2011 Sep; 22:1958-1972.

23. Peters U, Bien S, Zubair N. Genetic architecture of colorectal cancer. Gut. 2015 Oct;64(10):1623-36.

24. Hampel H, Pearlman R, Beightol M, Zhao W, Jones D, Frankel WL, Goodfellow PJ, Yilmaz A, Miller K, Bacher J, Jacobson A, Paskett E, Shields PG, Goldberg RM, de la Chapelle A, Shirts BH, Pritchard CC; Ohio Colorectal Cancer Prevention Initiative Study Group. Assessment of Tumor Sequencing as a Replacement for Lynch Syndrome Screening and Current Molecular Tests for Patients With Colorectal Cancer. JAMA Oncol. 2018 Jun 1; 4:806-813.

25. Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Working Group. Recommendations from the EGAPP Working Group: genetic testing strategies in newly diagnosed individuals with colorectal cancer aimed at reducing morbidity and mortality from Lynch syndrome in relatives. Genet Med. 2009 Jan; 11:35-41.

26. Li X, Liu G, Wu W. Recent advances in Lynch syndrome. Exp Hematol Oncol. 2021 Jun 12; 10:37.

27. Vasen HF, Möslein G, Alonso A, Aretz S, Bernstein I, Bertario L, Blanco I, Bülow S, Burn J, Capella G, Colas C, Engel C, Frayling I, Friedl W, Hes FJ, Hodgson S, Järvinen H, Mecklin JP, Møller P, Myrhøi T, Nagengast FM, Parc Y, Phillips R, Clark SK, de Leon MP, Renkonen-Sinisalo L, Sampson JR, Stormorken A, Tejpar S, Thomas HJ, Wijnen J. Guidelines for the clinical management of familial adenomatous polyposis (FAP). Gut. 2008 May; 57:704-13.

28. Win AK, Dowty JG, Cleary SP, Kim H, Buchanan DD, Young JP, Clendenning M, Rosty C, MacInnis RJ, Giles GG, Boussioutas A, Macrae FA, Parry S, Goldblatt J, Baron JA, Burnett T, Le Marchand L, Newcomb PA, Haile RW,

Hopper JL, Cotterchio M, Gallinger S, Lindor NM, Tucker KM, Winship IM, Jenkins MA. Risk of colorectal cancer for carriers of mutations in MUTYH, with and without a family history of cancer. Gastroenterology. 2014 May; 146:1208-11.e1-5.

29. Ma H, Brosens LAA, Offerhaus GJA, Giardiello FM, de Leng WWJ, Montgomery EA. Pathology and genetics of hereditary colorectal cancer. Pathology. 2018 Jan; 50:49-59.

30. Stoffel EM, Koeppe E, Everett J, Ulintz P, Kiel M, Osborne J, Williams L, Hanson K, Gruber SB, Rozek LS. Germline Genetic Features of Young Individuals With Colorectal Cancer. Gastroenterology. 2018 Mar;154: 897-905.

31. US Preventive Services Task Force, Bibbins-Domingo K, Grossman DC, Curry SJ, Davidson KW, Epling JW Jr, García FAR, Gillman MW, Harper DM, Kemper AR, Krist AH, Kurth AE, Landefeld CS, Mangione CM, Owens DK, Phillips WR, Phipps MG, Pignone MP, Siu AL. Screening for Colorectal Cancer: US Preventive Services Task Force Recommendation Statement. JAMA. 2017 Jun 6; 317:2239.

32. Ladabaum U, Dominitz JA, Kahi C, Schoen RE. Strategies for Colorectal Cancer Screening. Gastroenterology. 2020 Jan; 158:418-432.

33. Simon K. Colorectal cancer development and advances in screening. Clin Interv Aging. 2016 Jul 19; 11:967-76.

34. American Joint Committee on Cancer. Chapter 20 - Colon and Rectum. In: AJCC Cancer Staging Manual. 8th ed. New York, NY: Springer; 2017.

35. Grady WM, Markowitz SD. The molecular pathogenesis of colorectal cancer and its potential application to colorectal cancer screening. Dig Dis Sci. 2015 Mar; 60:762-72.

36. De Palma FDE, D'Argenio V, Pol J, Kroemer G, Maiuri MC, Salvatore F. The Molecular Hallmarks of the Serrated Pathway in Colorectal Cancer. Cancers. 2019 Jul 20; 11:1017.

37. Kasi A, Handa S, Bhatti S, Umar S, Bansal A, Sun W. Molecular Pathogenesis and Classification of Colorectal Carcinoma. Curr Colorectal Cancer Rep. 2020 Sep; 16:97-106.

38. Gupta R, Sinha S, Paul RN. The impact of microsatellite stability status in colorectal cancer. Curr Probl Cancer. 2018 Nov; 42:548-559.

39. Harada S, Morlote D. Molecular Pathology of Colorectal Cancer. Adv Anat Pathol. 2020 Jan; 27:20-26.

40. Nazemalhosseini Mojarad E, Kuppen PJ, Aghdaei HA, Zali MR. The CpG island methylator phenotype (CIMP) in colorectal cancer. Gastroenterol Hepatol Bed Bench. 2013.

41. Driehuis E, Kretzschmar K, Clevers H. Establishment of patient-derived cancer organoids for drug-screening applications. Nat Protoc. 2020 Oct; 15:3380-3409.

42. Schutgens F, Clevers H. Human Organoids: Tools for Understanding Biology and Treating Diseases. Annu Rev Pathol. 2020 Jan 24; 15:211-234.

43. Clevers H. Modeling Development and Disease with Organoids. Cell. 2016 Jun 16; 165:1586-1597.

44. Rookmaaker MB, Schutgens F, Verhaar MC, Clevers H. Development and application of human adult stem or progenitor cell organoids. Nat Rev Nephrol. 2015 Sep; 11:546-54.

45. Drost J, Clevers H. Translational applications of adult stem cell-derived organoids. Development. 2017 Mar 15; 144:968-975.

46. Lancaster MA, Knoblich JA. Organogenesis in a dish: modeling development and disease using organoid technologies. Science. 2014 Jul 18; 345:1247125.

47. Sato T, Vries RG, Snippert HJ, van de Wetering M, Barker N, Stange DE, van Es JH, Abo A, Kujala P, Peters PJ, Clevers H. Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. Nature. 2009 May 14; 459:262-5.

48. Sato T, Stange DE, Ferrante M, Vries RG, Van Es JH, Van den Brink S, Van Houdt WJ, Pronk A, Van Gorp J, Siersema PD, Clevers H. Long-term expansion of epithelial organoids from human colon, adenoma, adenocarcinoma, and Barrett's epithelium. Gastroenterology. 2011 Nov; 141:1762-72.

49. Drost J, Karthaus WR, Gao D, Driehuis E, Sawyers CL, Chen Y, Clevers H. Organoid culture systems for prostate epithelial and cancer tissue. Nat Protoc. 2016 Feb; 11:347-58.

50. Wang K, Yuen ST, Xu J, Lee SP, Yan HH, Shi ST, Siu HC, Deng S, Chu KM, Law S, Chan KH, Chan AS, Tsui WY, Ho SL, Chan AK, Man JL, Foglizzo V, Ng MK, Chan AS, Ching YP, Cheng GH, Xie T, Fernandez J, Li VS, Clevers H, Rejto PA, Mao M, Leung SY. Whole-genome sequencing and comprehensive molecular profiling identify new driver mutations in gastric cancer. Nat Genet. 2014 Jun; 46:573-82.

51. Sachs N, de Ligt J, Kopper O, Gogola E, Bounova G, Weeber F, Balgobind AV, Wind K, Gracanin A, Begthel H, Korving J, van Boxtel R, Duarte AA, Lelieveld D, van Hoeck A, Ernst RF, Blokzijl F, Nijman IJ, Hoogstraat M, van de

Ven M, Egan DA, Zinzalla V, Moll J, Boj SF, Voest EE, Wessels L, van Diest PJ, Rottenberg S, Vries RGJ, Cuppen E, Clevers H. A Living Biobank of Breast Cancer Organoids Captures Disease Heterogeneity. Cell. 2018 Jan 11; 172:373-386.e10.

52. Boj SF, Hwang CI, Baker LA, Chio II, Engle DD, Corbo V, Jager M, Ponz-Sarvise M, Tiriac H, Spector MS, Gracanin A, Oni T, Yu KH, van Boxtel R, Huch M, Rivera KD, Wilson JP, Feigin ME, Öhlund D, Handly-Santana A, Ardito-Abraham CM, Ludwig M, Elyada E, Alagesan B, Biffi G, Yordanov GN, Delcuze B, Creighton B, Wright K, Park Y, Morsink FH, Molenaar IQ, Borel Rinkes IH, Cuppen E, Hao Y, Jin Y, Nijman IJ, Iacobuzio-Donahue C, Leach SD, Pappin DJ, Hammell M, Klimstra DS, Basturk O, Hruban RH, Offerhaus GJ, Vries RG, Clevers H, Tuveson DA. Organoid models of human and mouse ductal pancreatic cancer. Cell. 2015 Jan 15; 160:324-38.

53. Broutier L, Mastrogiovanni G, Verstegen MM, Francies HE, Gavarró LM, Bradshaw CR, Allen GE, Arnes-Benito R, Sidorova O, Gaspersz MP, Georgakopoulos N, Koo BK, Dietmann S, Davies SE, Praseedom RK, Lieshout R, IJzermans JNM, Wigmore SJ, Saeb-Parsy K, Garnett MJ, van der Laan LJ, Huch M. Human primary liver cancer-derived organoid cultures for disease modeling and drug screening. Nat Med. 2017 Dec; 23:1424-1435.

54. Sachs N, Papaspyropoulos A, Zomer-van Ommen DD, Heo I, Böttinger L, Klay D, Weeber F, Huelsz-Prince G, Iakobachvili N, Amatngalim GD, de Ligt J, van Hoeck A, Proost N, Viveen MC, Lyubimova A, Teeven L, Derakhshan S, Korving J, Begthel H, Dekkers JF, Kumawat K, Ramos E, van Oosterhout MF, Offerhaus GJ, Wiener DJ, Olimpio EP, Dijkstra KK, Smit EF, van der Linden M, Jaksani S, van de Ven M, Jonkers J, Rios AC, Voest EE, van Moorsel CH, van der Ent CK, Cuppen E,

van Oudenaarden A, Coenjaerts FE, Meyaard L, Bont LJ, Peters PJ, Tans SJ, van Zon JS, Boj SF, Vries RG, Beekman JM, Clevers H. Long-term expanding human airway organoids for disease modeling. EMBO J. 2019 Feb 15;38.

55. Schutgens F, Rookmaaker MB, Margaritis T, Rios A, Ammerlaan C, Jansen J, Gijzen L, Vormann M, Vonk A, Viveen M, Yengej FY, Derakhshan S, de Winter-de Groot KM, Artegiani B, van Boxtel R, Cuppen E, Hendrickx APA, van den Heuvel-Eibrink MM, Heitzer E, Lanz H, Beekman J, Murk JL, Masereeuw R, Holstege F, Drost J, Verhaar MC, Clevers H. Tubuloids derived from human adult kidney and urine for personalized disease modeling. Nat Biotechnol. 2019 Mar; 37:303-313.

56. Lancaster MA, Renner M, Martin CA, Wenzel D, Bicknell LS, Hurles ME, Homfray T, Penninger JM, Jackson AP, Knoblich JA. Cerebral organoids model human brain development and microcephaly. Nature. 2013 Sep 19; 501:373-9.

57. Xie BY, Wu AW. Organoid Culture of Isolated Cells from Patient-derived Tissues with Colorectal Cancer. Chin Med J . 2016 Oct 20.

58. Yang L, Yang S, Li X, Li B, Li Y, Zhang X, Ma Y, Peng X, Jin H, Fan Q, Wei S, Liu J, Li H. Tumor organoids: From inception to future in cancer research. Cancer Lett. 2019 Jul 10.

59. Ji DB, Wu AW. Organoid in colorectal cancer: progress and challenges. Chin Med J. 2020 Aug 20.

60. Vlachogiannis G, Hedayat S, Vatsiou A, Jamin Y, Fernández-Mateos J, Khan K, Lampis A, Eason K, Huntingford I, Burke R, Rata M, Koh DM, Tunariu N, Collins D, Hulkki-Wilson S, Ragulan C, Spiteri I, Moorcraft SY, Chau I, Rao S, Watkins D, Fotiadis N, Bali M, Darvish-Damavandi M, Lote H, Eltahir Z, Smyth

EC, Begum R, Clarke PA, Hahne JC, Dowsett M, de Bono J, Workman P, Sadanandam A, Fassan M, Sansom OJ, Eccles S, Starling N, Braconi C, Sottoriva A, Robinson SP, Cunningham D, Valeri N. Patient-derived organoids model treatment response of metastatic gastrointestinal cancers. Science. 2018 Feb 23; 359: 920-926.

61. Fernando TM, Piskol R, Bainer R, Sokol ES, Trabucco SE, Zhang Q, Trinh H, Maund S, Kschonsak M, Chaudhuri S, Modrusan Z, Januario T, Yauch RL. Functional characterization of SMARCA4 variants identified by targeted exome-sequencing of 131,668 cancer patients. Nat Commun. 2020 Nov 3; 11:5551.

62. Ogunwobi OO, Mahmood F, Akingboye A. Biomarkers in Colorectal Cancer: Current Research and Future Prospects. Int J Mol Sci. 2020 Jul 27; 21:5311.

63. Margonis GA, Spolverato G, Kim Y, Karagkounis G, Choti MA, Pawlik TM. Effect of KRAS Mutation on Long-Term Outcomes of Patients Undergoing Hepatic Resection for Colorectal Liver Metastases. Ann Surg Oncol. 2015 Dec; 22: 4158-65.

64. Raskov H, Søby JH, Troelsen J, Bojesen RD, Gögenur I. Driver Gene Mutations and Epigenetics in Colorectal Cancer. Ann Surg. 2020 Jan; 271:75-85.

65. Venderbosch S, Nagtegaal ID, Maughan TS, Smith CG, Cheadle JP, Fisher D, Kaplan R, Quirke P, Seymour MT, Richman SD, Meijer GA, Ylstra B, Heideman DA, de Haan AF, Punt CJ, Koopman M. Mismatch repair status and BRAF mutation status in metastatic colorectal cancer patients: a pooled analysis of the CAIRO, CAIRO2, COIN, and FOCUS studies. Clin Cancer Res. 2014 Oct 15; 20: 5322-30.

66. Yokota T. Are KRAS/BRAF mutations potent prognostic and/or predictive biomarkers in colorectal cancers? Anticancer Agents Med Chem. 2012 Feb; 12:163-71.