UNIVERSITA' DEGLI STUDI DI PADOVA

CORSO DI LAUREA MAGISTRALE IN MEDICINA E CHIRURGIA

Dipartimento di Scienze Chirurgiche Oncologiche e Gastroenterologiche
Direttore: Prof. Salvatore Pucciarelli

TESI DI LAUREA:

PATHOLOGICAL AND MOLECULAR FEATURES OF SPORADIC COLON CANCER IN YOUNG ADULTS

Relatore:
Prof.ssa Spolverato Gaya

Laureanda:
Vagnarelli Giulia

ANNO ACCADEMICO 2021/22

UNIVERSITA' DEGLI STUDI DI PADOVA

CORSO DI LAUREA MAGISTRALE IN MEDICINA E CHIRURGIA

Dipartimento di Scienze Chirurgiche Oncologiche e Gastroenterologiche
Direttore: Prof. Salvatore Pucciarelli

TESI DI LAUREA:

PATHOLOGICAL AND MOLECULAR FEATURES OF SPORADIC COLON CANCER IN YOUNG ADULTS

Relatore:
Prof.ssa Spolverato Gaya

Laureanda:
Vagnarelli Giulia

ANNO ACCADEMICO 2021/22

Summary

Riassunto

Presupposti dello studio

In tutto il mondo, il cancro al colo-retto nei soggetti giovani adulti (<50 anni) sta registrando un allarmante aumento dei casi diagnosticati e della mortalità cancro associata. Per quanto nel cancro al colon-retto ad esordio precoce la predisposizione genetica rivesta un ruolo centrale, ad oggi la maggioranza dei casi rimane sporadica. Ciononostante, l'intero spettro di sequenze mutazionali germinali e somatiche implicate nel processo rimane ancora sconosciuto.

Scopo dello Studio

Nonostante l'aumento generale dei casi registrati di cancro al colon-retto tra i soggetti giovani adulti in tutto il mondo, nessuna effettiva spiegazione è ancora stata individuata. Lo scopo dello studio è di indagare il microambiente immunitario tumorale, le mutazioni onogeniche e l'instabilità microsatellitare rilevabili in pazienti giovani (≤50anni) con cancro al colon, col fine di meglio definire l'origine molecolare e citologica di quest'ultimo.

Metodi

Lo studio condotto ha caratteristiche retrospettive e osservazionali. La prima parte dell'analisi ha incluso pazienti affetti da cancro al colon e operati da Gennaio 2015 a Marzo 2022 presso la Clinica Chirurgica 3 dell'Azienda Ospedale Università di Padova. L'intento primario dello studio riguardava le differenze nel microambiente tumorale tra il cancro al colon in pazienti giovani e anziani, considerando 50 anni come limite di età predefinito. In particolare, un totale di 502 pazienti è stato reclutato, e una analisi sull'infiltrato istologico a livello linfomononucleare, vascolare e perineurale, così come la valutazione dell'immunoistochimica per MLH1, PMS2, MSH2 e MSH6 e i profili mutazionali dei geni KRAS, BRAF, NRAS sono stati eseguiti. Al fine di ampliare la prospettiva sulle mutazioni genetiche nel cancro al colon e per investigare più approfonditamente il coinvolgimento genetico del sistema immunitario nel suo sviluppo, un ulteriore insieme di 378 pazienti con cancro al colon sono stati analizzati utilizzando il database pubblico Cancer Genome Atlas Research Network (*TCGA, PanCancer Atlas – Colorectal Adenocarcinoma*). Il profilo mutazionale per i geni APC, TP53, TTN, KRAS, MUC16, PIK3CA, CSMD3, SYNE1, RBFOX1, CSMD1, l'analisi delle mutazioni genetiche più frequenti con

significatività statistica (ADM2, LDLRAD4, SIN3A, ANXA7, EGFL6, SCL264, BTN2A2, DNAJB9, ZFN304, PPP6C) e l'espressione nucleare dei più comuni marcatori linfocitari sono stati eseguiti ponendo a confronto l'età dei pazienti giovani (≤50 years) e vecchi (>50 anni) alla diagnosi.

Risultati

I differenti profili oncogenetici, microsatellitari e infiammatorio tumorale sono stati analizzati. Per i primi, i risultati hanno mostrato una maggiore profilo mutazionale di NRAS tra i casi di cancro ad esordio giovanile (p=0.031), al contrario di BRAF, che ha mostrato una prevalenza significativa di espressione nei soggetti anziani (p=0.069). A riguardo dei microsatelliti e dell'infiltrato infiammatorio tumorale, l'analisi ha mostrato in entrambi i gruppi un simile andamento, risultando un tasso inferiore di coinvolgimento microsatellitare con alta probabilità nei tumori ad insorgenza giovanile (mostrando in particolare una mancanza totale di assenza di espressione di MSH2), insieme ad una più spiccata assenza dell'infiltrato tumorale nelle sue tre forme (linfomononucleare, vascolare e perineurale) nello stesso gruppo di pazienti. Analizzando le mutazioni genetiche più frequenti, i geni MUC16, CSMD1 e BTN2A2 sono risultati essere i geni coinvolti nella risposta immunitaria più frequentemente alterati nei giovani adulti. Infine, l'analisi dell'espressione nucleare dei marcatori linfocitari più comuni ha mostrato nei giovani (30-50 anni alla diagnosi) un andamento generale di diminuita espressione riguardante i geni GZMB, HLA-B e HLA-C, e di contemporanea aumentata espressione di IL17B.

Conclusioni

Tra i giovani adulti il cancro al colon-retto sta registrando un allarmante aumento dell'incidenza in tutto il mondo, senza che alcuna motivazione scientifica sia stata individuata finora a riguardo. Lo studio ha rilevato una riduzione generale dell'espressione di linfociti citotossici e regolatori in pazienti giovani adulti, suggerendo che parte dei casi di cancro al colon ad esordio giovanile possa essere dovuta ad una debolezza nell'immunosorveglianza di questi soggetti. Sebbene ulteriori analisi a riguardo siano necessarie, i risultati della ricerca operata in questo studio, focalizzandosi sugli aspetti più cruciali della genesi del cancro al colon, si propongono come una guida per una più approfondita condiscienza di questa realtà sempre più problematica.

Abstract

Background

All over the world, early-onset colorectal cancer is registering an alarming increase in diagnosed cases among young subjects (<50 years old) together with an increase of CCR-related mortality among younger patients over the past few decades. Although in early onset colon cancer genetic predisposition is shown to have a central role, most cases still remain sporadic. Even so, the full spectrum of germline and somatic sequence variations implicated are unknown. Different risk factors have been evaluated: changing in diet habits and overweight together with smoking, sedentary behaviour and antibiotic use have been found related to an increase of young-onset CCR. Latest theories present Exposome as a new possible aetiology factor of sporadic young-onset CCR.

Aim of the study

As colorectal cancer among young subjects is registering a worldwide increase, no actual explanation has been found so far. The aim of this study was to assess tumoral mucosa immune microenvironment, microsatellite instability and oncogene mutation of colon cancer in young patients (≤50 years), in order to better define its molecular and cytological genesis.

Methods

A retrospective observational study was conducted. The first part of the study included patients who were under colon surgery from January 2015 to March 2022 at the General Surgery 3 Unit of the Azienda Ospedale Università di Padova. The primary focus was on cytological and molecular differences between colon cancer in young and elderly patients, using 50 years as a predefined age cut-off. Histological lymphomonocytic, vascular and perinuclear infiltration were performed, as well as immunochemistry for MLH1, PMS2, MSH2, and MSH6 and hotspot mutational profile of KRAS, BRAF, NRAS genes. A total of 502 patients was recruited.

In order to have a wider perspective on gene mutational colon cancer profile and better investigate immunity system genes involvement in CCR developing, a total of 378 colon cancer patients were also examined using public database Cancer Genome Atlas Research Network (*TCGA, PanCancer Atlas - Colorectal Adenocarcinoma*). APC, TP53, TTN, KRAS, MUC16, PIK3CA, CSMD3,

SYNE1, RBFOX1, CSMD1 genes mutational profile were analysed together with the most frequently significant p-values in gene alterations (ADM2, LDLRAD4, SIN3A, ANXA7, EGFL6, SCL264, BTN2A2, DNAJB9, ZFN304, PPP6C), comparing early (≤50 years) and elderly (>50 years) patient's age at the diagnosis. Also, mRNA expression of most common lymphocyte markers was assessed.

Results

The different expression of oncogenes, microsatellite pattern and peritumoral infiltrate were analysed. As for the oncogenes results, NRAS mutation was mainly detected in young-onset cancer profile rather than in old-one (p=0.031), on the contrary of BRAF, showing significant prevalence in old subjects (p=0.069). KRAS mutations did not show any significant difference between the two groups. Microsatellites and tumoral infiltrate showed similar patterns in the two clusters, presenting the young-onset group a lower rate of high probability microsatellite involvement (especially displaying a total absence of MSH2 expression) and a predominant absence of the tumoral infiltrate pattern in all its three forms (lymphomononuclear, vascular and perineural). The rate of the most frequently mutated genes showed a mutation of MUC16, CSMD1 and BTN2A2 immune genes prevalence in young subjects comparing to elderly ones (>50 years). Among mRNA expression of most common lymphocyte markers, GZMB, HLA-B and HLA-C in particular showed a general trend of reduced amount in young subjects (30-50 years at the diagnosis), while IL17B was found to be increased in the same group of patients.

Conclusions

Colorectal cancer among young subjects is registering a worldwide increase without any explanation found so far. An overall reduced immune expression of cytotoxic and regulatory lymphocytes in young patients was found, suggesting part of early onset colon cancer might be due to a weak immunesurveillance. Although further studies are needed, these findings, enlightening the crucial aspects of colon cancer genesis, lead the way to a better comprehension to this always more problematic reality.

Introduction

Colorectal cancer is the second most frequently diagnosed cancer by incidence among all malignancies and the second leading cause of cancer-related deaths [1] [2]. Comprehensively the third most frequently diagnosed cancer in men and second in women [1], colorectal cancer responds to 10% overall diagnosed cancer per year [1] (1.8 million new cases and about 881,000 deaths worldwide in 2018 [3]).

All over the world, early-onset colorectal cancer is registering an alarming increase in diagnosed cases among young subjects (<50 years old) together with an increase of CCR-related mortality among younger patients (10% of all new diagnoses of this cancer [4]) over the past few decades, since the 1980s [5] [6] [2]. In addition, this trend has been shown to be in opposite contrast with the steady decline in the incidence and mortality of colorectal cancer in later-onsets. According to these changes, the medium age at the diagnosis shifted from 72 to 66 years, and in the next 10 years 12% of colon cancer will be diagnosed in younger than 50 years patients [4][7][8].

Epidemiology

CCR is primarily an older adult disease, with a median age of 65 years [9]. The incidence of the cancer in old population (>50 years) is currently decreasing and an improved survival of older age patients has been registered, thanks to the improve use of screenings, better therapies and supportive cares. On the other hand, young patients are found to have an always higher incidence of CCR, maybe due to low suspicious together with lifestyle and dietary changes [1].

Aetiopathogenesis

For the aetiopathogenesis, 70% of all cases of colorectal cancer are sporadic [9], while the remaining 30% are to be attributed to specific gene mutations, 5% of which are caused by hereditary syndromes [1]. Currently, even though young subjects with CCR are more likely expected to have an hereditary syndrome comparing to elderly people, the majority of cases remain sporadic, with no identifiable cause [4].

Pathogenesis

For what concerns colorectal cancer pathogenesis, it has been found to respond to three different neoplastic pathways [1]. The first one is the adenoma-carcinoma pathway, which represents the 70-90% of cases [1]. In particular, FAP and sporadic cancer are based on this pathway. The process takes origin from an aberrant crypt, that will evolve into a neoplastic precursor lesion (represented by the polyp), which would eventually turn into colorectal cancer in about 10-15 years. The pathway is due to an accumulation of different genetic mutations (APC, KRAS, BRAF, SMAD4 and TP53) deriving from chromosomal instability (CIN). The second pathway regards serrated (traditional and sessile) neoplasia, accounting for about 10-20% of cases [1]. In this case the process is based on genetic (mainly KRAS and BRAF, followed by further genes) and epigenetic mutations (CIMP, CpG island methylator phenotype), which lead to microsatellite instability in the end (eg. MLH1) [1]. Finally, the third pathway principally regards Lynch syndrome, leaving the 10-15% of all these cases to other forms of colorectal cancer [1]. It is caused by microsatellite instability, in particular by germline mutations in MMR genes (eg. MSH2,6, MLH1,3, PMS2,6) [1].

Yong-onset and late-onset comparison

Sporadic tumors can be defined as an elderly-onset tumor, as it mainly occurs at the age of 70-75 years [10]. Young age at the CCR diagnosis moment instead, might suggests an hereditary syndrome basis [11] [12], even though only the 5-7% are caused by a well defined germline mutation in a cancer instability gene [1]. In particular, young-onset CCR are principally attributed to Lynch Syndromes, FAP and MUTYH-associated polyposes (MAP) [7]. Comparing young and late-onset CCRs, they have been found to have histological significant differences in locational site, stage of presentation and molecular profile [11]. Specifically, young-onset colorectal tumor principally occurs in distal colon and rectum, and presents a higher percentage of synchronous and metachronous tumors. Moreover, a higher incidence of recurrence and metastasis development, a higher number of mucinous and signet ring features, and a poorly differentiated histology have been found in young-onset cancers [11]. As for the molecular features, the majority of young-onset CCR shows MSS (microsatellite stability) and lack of DNA repair mechanism abnormalities, on the contrary of late-onset cases [11] [13]. In particular, MSI in younger patients can be more frequently

linked to Lynch Syndrome and less to epigenetic inactivation of MLH1 - as well as MSH2 inactivation - which is preferentially shown in late-onset colorectal cancer, together with MLH1 promoter methylation [11]. Also, CpG island methylator phenotype (CIMP) mutations are found to be low in young-onset cancers [11] [14]. MSS tumors on the other hand show different locational sides between young (left) and late (right) onsets, while MACS (microsatellite and chromosome stable CCR) can be most frequently found in younger cases and are mainly related to worse prognosis [11]. Finally, LINE1-hypomethylation, linked to a higher colorectal cancer mortality, is less shown in young-onset colon cancer [11].

Tumoral mucosa immune microenvironment

This study finds one of its key points in the tumoral mucosa immune microenvironment of colorectal cancer. Tumour-infiltrating immune cells have shown to be of primary importance to determine the outcome of the patients, both in disease-free survival and overall survival [10]. In particular, histological datas of early metastatic invasion - vascular emboli, lymphatic invasion and perineural invasion (VELIPI) – has been found to significantly influence disease-free and overall survival [10]. In particular, the inflammatory microenvironment of VELIPI-negative tumours was found to have a higher density of effector memory T cells (positive markers: CD3,CD8,CD45R0,CDR7,CD27,CD28) together with a higher density of CD45R0+ cells, that correlates with a good clinical outcome [10]. Also, the meaning of Th1 effector cells and related cytokines as a favourable prognostic sign was demonstrated by the higher tumoral mRNA and markers (CD8, T-BET, IFNregultoryfator1, IFNγ, granulosin, granzymB) levels [10]. VELIPI-negative tumours are also associated with both a higher density of T memory cells of all stages of differentiation and of mature T cells (low density shown in N2-N3 staged tumours), correlating to longer disease-free and overall survival [10].

Symptoms

Even though colorectal cancer remains principally an asymptomatic disease, the most common symptoms at the presentation are rectal bleeding (in both benignant and malignant cancer forms), change in bowel habits, abdominal pain and new anaemia [1]. In younger subjects other supplementary factors that can be

considered are family history for CCR, positive fecal occult blood (FOB) or melena, inexplicable weight loss [1].

Diagnosis

As for the diagnosis of colorectal cancer, it finds its gold standard in endoscopy, both a diagnostic and potentially treating method [1]. Imaging can also help in staging the tumour, and is mainly used for locoregional and metastatic staging [1]. Also, determination of CEA has been found to be a useful prognostic marker, especially in postoperative patients [1]. TNM staging system, histological datas (lymphatic, perineural and venous invasion), immunoscore and mismatch-repair testing have all become routine examinations to outline the cancer and the best treatment for it [1].

Treatment

So far, surgery remains the best option for curative purpose treatment, while radiotherapy is mainly used for rectal cancers [1]. As a local treatment can be cited both resection or local ablation for liver metastases, stereotaxic radiotherapy and radiofrequency ablation for lung metastases.

As for systemic treatment, traditional chemotherapy, biological therapy, savage therapy drugs, targeted therapies and immunotherapy can all be used in various combinations. In particular, as a first line therapy, Fluoracil (Capecitabine) has been demonstrated to be very effective [1]. Moreover, lately it has been adopted the addition of Oxliplatin or Irinotecan as a new standard, in order to reduce side effects such as cumulative sensory neuropathy [1]. For about metastatic cancers, biological treatment can be used, as well as salvage therapy (eg. Regorafeinb), especially for refractory metastatic cancer [1]. In particular, anti-VEGF and anti-EGFR drugs can be used for left-sided RAS and RAF-wt metaplasic colorectal cancer [1]. V600E BRAF mutation instead is linked to a 2-3 times worse outcome for its intrinsically resistance to anti-EGFR therapy [1]. Immunotherapy is mainly used for tumours with specific kind of dMMR or high MSI (eg. Nivolumab and Ipilimumab) [1]. New markers are emerging as researcher proceeds [1].

The issue and the exposome theory

Rising incidence of young-onset colorectal cancer is now one of the biggest epidemiological crises facing the world of surgical oncology [15]. Within the next decade, 1 out of 10 colon cancers are estimated to be diagnosed in young adults

(>50years) [16]. *The reasons for this increasing incidence are so far unknown* [16][1][2].

Although in early onset colon cancer genetic predisposition is shown to have a central role, most cases still remain sporadic [16]. Even so, the full spectrum of germline and somatic sequence variations implicated are unknown [16]. A new theory in sporadic young onset colorectal cancer aetiopathogenesis points the exposome as a possible cause factor.

The exposome represents the totality of a subject exposures from conception onwards and can be defined as the environmental equivalent of the human genome [16][17]. Gene and environment interactions (microbiome and germline genetics, GxE) are thought to possibly have important roles in the aetiology of early- onset colorectal cancer [2]. In particular, the exposome involves two different macro-areas: external environment (diet, smoking, alcohol, antibiotic exposure, in vitro fertilization, cesarian delivery), followed by internal environment (gut microbiota, IBD, chronic health conditions) [2][16].

As a matter of fact, clinical and pathological features in early and late-onset colon cancer are different, with the former presenting advanced stage at diagnosis and unfavourable histopathological qualities. According to Alexandra M. Zaborowski et al [16] standardized, age-specific preventive, screening, diagnostic, and therapeutic strategies are required to optimize outcomes.

Aim of the study

As colorectal cancer among young subject is registering a worldwide increase, no actual explanation has been found so far. The aim of this study was to assess tumoral mucosa immune microenvironment, microsatellite instability and oncogene mutation of colon cancer in young patients (≤50 years), in order to better define its molecular and cytological genesis.

Patients and methods

Study design (Azienda Ospedale Università di Padova cohort)

As a first research outline, clinical and pathological records of a series of consecutive colon cancer patients operated on at the General Surgery 3 Unit of the Azienda Ospedale Università di Padova from January 2015 to March 2022 were retrieved for this retrospective and observational study. We defined young colon cancer patients using 50 years as predefined age cut-off. Histology for the infiltration of intratumoral lymphomononuclear cells, immunohistochemistry for MLH1, PMS2, MSH2, and MSH6 to define microsatellite stability or instability (MSS/MSI), and mutational analysis of BRAF, KRAS, and NRAS were all performed. Young and old colon cancer patients were compared. All participants provided their informed consent.

Study design and and patients selection are shown in *Figure 1*.

Study design (TCGA cohort, PanCancer Atlas - Colorectal Adenocarcinoma)

A total of 378 colon cancer patients were also examined using public database Cancer Genome Atlas Research Network (TCGA, PanCancer Atlas - Colorectal Adenocarcinoma) [18]. Patients were selected among Colorectal Adenocarcinoma database for Colon Cancer cases. Rectal Adenocarcinoma and Mucinous Adenocarcinoma of the Colon and Rectum were excluded (see *Figure 2*). Original data were plotted on a novel database and analysed. Early (≤50 years) and elderly (>50 years) patient's age at the diagnosis were compared for highest average frequency cancer genes found mutated (APC, TP53, TTN, KRAS, MUC16, PIK3CA, CSMD3, SYNE1, RBFOX1, CSMD1). Furthermore, the most significant p-values in gene alterations were considered (ADM2, LDLRAD4, SIN3A, ANXA7, EGFL6, SCL264, BTN2A2, DNAJB9, ZFN304, PPP6C).

In the study, an analysis on mRNA expression of most common lymphocytes markers (CD4, CD8A, CD3, IL17B, T-bet (TBX21), GZMB, IFNB1, FOXP3, CD25, HLA-A, HLA-B, HLA-C, HLA-DR, CD40, CD80, CD86, PD1, PDL1, CTLA4) was displayed also, using TCGA Colon Cancer Cases database (panCnacer Atlas – 2018).

Study design and patients' selection are shown in *Figure 2*.

Figure 1. *Study design and patients' selection for AOPD (Azienda Ospedale Università di Padova). CC: Colon Cancer*

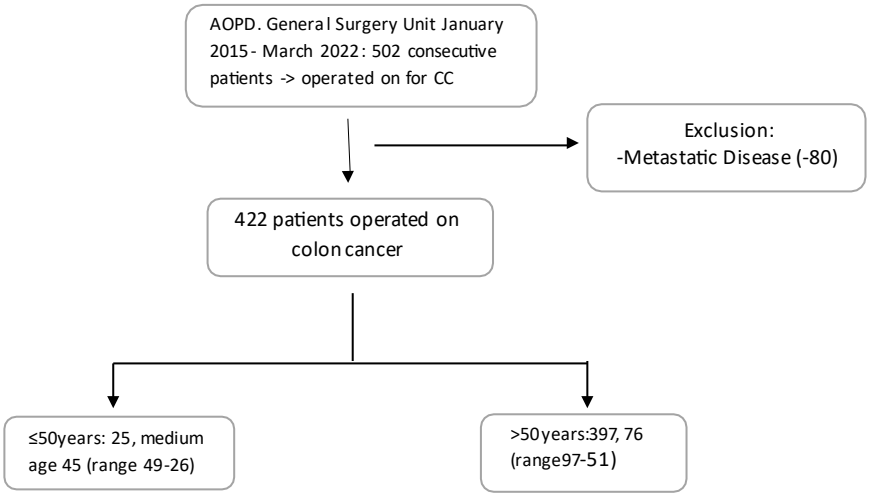
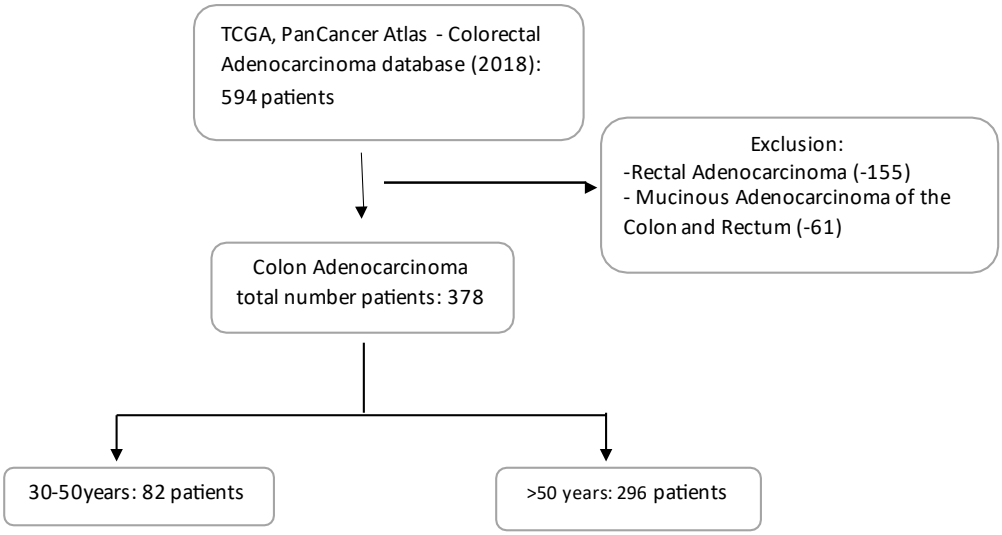


Figure 2. *Study design and patients' selection for TCGA, PanCancer Atlas - Colorectal Adenocarcinoma database (2018).*



Histopathology (AOPD)

The histopathological examination of all resected specimens consisted in the evaluation of the tumor stage, residual tumor, grading, and the number of lymph nodes involved. Specimens were fixed in 10% formaldehyde and set in paraffin.

The lymph nodes were counted and assessed by a pathologist. The nodal status (N0, N1) was evaluated in accordance with the eighth edition of the TNM classification [19]. Furthermore, the infiltration of lymphomononuclear cells was graded as absent, low grade, or high grade. The number of tumor infiltrating lymphocytes (TILS) was also obtained by considering the mean number of positive cells observed under a 5 high power field (40).

Immunohistochemistry (AOPD)

Immunohistochemical (IHC) analyses were performed using standard procedures, and the resulting sections were evaluated by a single pathologist in a blinded fashion. Immunocomplexes were detected using the Dako Real Envision System Peroxidase and 3- 3'di-aminobenzidine tetrahydrochloride chromogen as a substrate (Dako, Glostrup, Denmark) in formalin-fixed paraffin-embedded sections. IHC staining was performed using monoclonal antibodies for MLH1 (clone ES05, 1:100; Dako, Glostrup, Denmark), PMS2 (clone EP51, 1:100; Dako, Glostrup, Denmark), MSH2 (clone FE11, 1:100; Dako, Glostrup, Denmark), MSH6 (clone EP49, 1:100; Dako, Glostrup, Denmark), PD-L1 (clone 22C3, 1:50; Dako, Glostrup, Denmark), CD80 (clone 37711, 1:40; R & D Systems, Inc.), and CD8 (clone C8/144B, 1:200; Dako, Glostrup, Denmark). Sections were lightly counterstained with hematoxylin and performed automatically. Furthermore, for all the immunohistochemical markers (CD4, CD8, Tbet [T-bet: T-box expressed in T cells, a Th1 marker], FoxP3 [forkhead box P3, a Treg marker], PD1 [programmed death 1], PDL1 [programmed death ligand 1], and CD80) the absolute number of positive cells was obtained by considering the mean number of positive cells observed under a 5 high power field (40x). The Tbet/ CD4 and FoxP3/CD4 ratios were calculated, with a high ratio being defined as a ratio over 1. To assess the Th1/Treg ratio, frequencies of patients with a high Tbet/CD4 and FoxP3/CD4 ratio were compared with those of patients with a low ratio.

Analysis of MSI (AOPD)

DNA mismatch repair machinery-deficient tumors (MMRd) were defined by the absence of nuclear staining in one of the MLH1/PMS2 or MSH2/MSH6 pairs in tumor cells, as assessed in the colorectal setting [20]. The normal staining pattern of MLH1, PMS2, MSH2, and MSH6 was nuclear and was defined as MMR

proficient (MMRp). Infiltrating lymphocytes and stromal cells served as the internal positive controls. Microsatellite instability was studied using the following markers (BAT25, BAT26, D2S123, D17S250, D5S346, NR21, NR24, D18S58, BAT40, TGFbRII, TPOX, and TH01) using the Titano kit (Diatech Pharmacogenetics, Jesi, Italy) according to the Bethesda panel proposed by Bocker et al [21]. Extracted DNA (5 mL; ~20 ng) were used in a 50 mL PCR reaction, which contained 10X buffer, MgCl2 (1.5 mmol/L), dNTPS (200 mm), primers (5 mm), and dH2O. The PCR was performed on a GeneAmp PCR system 2720 Thermal Cycler (Applied Biosystems, Foster City, CA) using the following PCR cycling conditions: initial denaturation (8 minutes at 95°C), followed by 10 cycles denaturation (30 seconds at 94°C), annealing (45 seconds at 60°C), and extension (30 seconds at 72°C), followed by 22 cycles denaturation (30 seconds at 92°C), annealing (45 seconds at 55°C), and extension (30 seconds at 72°C). There was a final step of 10 minutes at 72°C and then held at 4°C. The PCR products were analysed on a 10% to 12% nondenaturing polyacrylamide gel and stained by Silver stain (Bio-Rad, San Diego, CA). The samples were run on an ABI PRISM 3130 Genetic Analyzer (Applied Biosystems, Foster City, CA) and analysed using GeneScan 3.7 software (Applied Biosystems, Foster City, CA). MSI was defined as the presence of additional bands in the PCR amplified product derived from neoplastic lesions in comparison to nonneoplastic tissues from the same patient.

BRAF, KRAS, and NRAS molecular profiling (AOPD)

The BRAF, KRAS, and NRAS status was obtained from diagnostic surveys (Sequenom MassArray and Sanger sequencing). DNA was extracted from the formalin-fixed paraffin-embedded samples after enrichment of neoplastic cellularity (ie, at least 25% of neoplastic cells in the sample). Serially cut 5-mm-thin sections were set on uncharged slides, deparaffinized, and lightly counterstained with hematoxylin. Microdissection was manually performed (under a light microscope) using a sterile injection needle. DNA was extracted using the QIAamp DNA FFPE tissue kit (Qiagen, Milan, Italy) and qualified as described previously [22].

Statistics (AOPD)

Continuous data were expressed as median and interquartile range, and categorical data as number and percentage. All tests were 2-sided. Statistical analysis was performed using R 3.3 (R Foundation for Statistical Computing, Vienna, Austria) [23].

Mutational Analysis (TCGA)

TCGA sequence information was obtained from the database of Genotypes and Phenotypes (dbGaP). Data from paired tumor and germline samples were independently aligned to human reference GRCh37-lite using BWA (Li and Durbin, 2009) v0.5.9 and de-duplicated using Picard 1.29. GenomeVIP (Mashl et al., 2017) was used to orchestrate germline calling using the following tools. Germline single nucleotide variants (SNVs) were identified using Varscan (Koboldt et al., 2012) version 2.3.8 (default parameters, except where –min-var-freq 0.10,–p value 0.10,–min-coverage 3,–strand-filter 1) operating on an mpileup stream produced by samtools (Li et al., 2009) version 1.2 (default parameters, except where -q 1 -Q 13) and GATK (McKenna et al., 2010) version 3.5 using the haplotype caller in single-sample mode with duplicate or unmapped reads removed and calls with quality threshold of 10 retained. Germline indels were identified using Varscan and GATK, both as configured as above, along with Pindel (Ye et al., 2016) version 0.2.5b8. We specified an insert size of 500 whenever this information was not present in the BAM header. Variants were limited to coding regions of full length transcripts obtained from Ensembl release 70 plus two additional base pairs flanking each exon that cover splice donor/acceptor sites. The union of GATK and VarScan SNVs was processed through our in-house false-positive filter (Kanchi et al., 2014). We included indels called by at least two out of the three callers (GATK, Varscan, Pindel) and high-confidence, Pindel-unique calls (at least 30x coverage and 20% VAF). The combined indels set was again processed through our false-positive filter (default parameters), except where-min-homopolymer 10–min-var-freq 0.2–min-var-count = 6. For germline and somatic variant comparison we restricted our data to the overlap of samples with at least one mutation in the MC3 MAF after restricting variants as outlied below. This overlap removed one gene from the germline predisposition list (CYLD). [24]

A publicly available MAF file was compiled by the TCGA MC3 Working Group and annotated with filter flags to highlight potential artifacts and discrepancies (Ellrott et al., 2018). A host of possible artifacts were flagged, including strand-bias, contamination, Oxo-guanine artifacts, and low normal read depth. If a mutation escaped flagging and was called by 2 or more variant calling tools, it was labeled a 'PASS'. We restricted analysis to PASS calls, except for samples from OV and LAML, which were early entrants in TCGA that were whole genome amplified (WGA). Of the 412 OV and 141 LAML samples in our dataset, 347 (84%) and 141 (100%), respectively, had artificial variants induced by WGA. In order to maintain sample sizes and uniformity in mutation calling, we did not filter mutations containing only 'wga' filter tags from these two cancer types. Seven bioinformatic tools were applied, five for Single Nucleotide Variants (SNV) and three for short Insertion Deletion (INDEL) events, with Varscan 2 providing both types of analysis. This list is comprised of MuTect (Cibulskis et al., 2013), VarScan2 (Koboldt et al., 2012), Indelocator (Chapman et al., 2011), Pindel (Ye et al., 2016), SomaticSniper (Larson et al., 2012), RADIA (Radenbaugh et al., 2014), and MuSE (Fan et al., 2016). The final call set was filtered to identify cohort level artifacts and was subject to extensive variant, subject, and cohort level QC. In total, 22,485,627 putative variants were identified and 2,907,335 high confidence mutations were retained after filtering [*24*][18].

Expressional Analysis

Total RNA for each sample was converted into a library of template molecules for sequencing on the Illumina Cluster Station and Genome Analyzer according to the protocol for the Illumina mRNA Sample preparation kit (Part#1004898, Rev A: Illumina, San Diego, CA). Briefly, poly-A mRNA was purified from total RNA (2 μg) using poly-T oligo-attached magnetic beads. The mRNA was then fragmented, and the first strand of cDNA was synthesized from the cleaved RNA fragments using reverse transcriptase and random primers. Following the synthesis of the second strand of cDNA, end repair was performed on overhangs using T4 DNA polymerase and Klenow DNA polymerase, followed by ligation of sequencing Adapters to the ends of the DNA fragments. The cDNA fragments were purified using a gel run at 80 V for approximately 3 hours until the Orange G dye band reached the bottom of the gel. The gel was stained with SYBR green

to visualize the DNA band. A band at 350 – 450 bp was excised vertically from the gel, which was then dissolved at room temperature using a QIAquick Gel Extraction Kit (Qiagen, Valencia, CA). The purified cDNA templates were enriched for 15 cycles of PCR amplification and validated using a BioAnalyzer to assess size, purity and concentration of the purified cDNA libraries. The cDNA libraries were placed on an Illumina Cluster Station for single end cluster generation according to the protocol outlined in the Illumina Genome Analysis User Guide (Part# 11251649, RevA). The template cDNA libraries (1.5 μg) were hybridized to a flow cell, amplified and linearized and denatured to create a flow cell with ssDNA ready for sequencing. Each flow cell was sequenced on an Illumina GAIIX Genome Analyzer. Each sample underwent a single lane of sequencing using single end sequencing for 76 cycles according to the protocol outlined in the Illumina Genome Analysis User Guide (Part# 11251649, RevA). After completion of the 76-cycle sequencing run, the raw sequence data entered the UNC RNAseq Workflow [*25*][18].

Results

Study patient characteristics

A total of 502 patients that went under surgery at General Surgery 3 Unit of the Azienda Ospedale Università di Padova from January 2015 to March 2022 were retrieved. To distinguish early-onset from late-onset Colon Cancer, 50 years was used as predefined age cut-off. Patients' characteristics and molecular and immunological data are shown in *Table 1*.

Table 1

Features			≤50 years	>50 years	P value
Total patients			28	474	
Age, mean (range)			45(49-26)	76(97-51)	
Cancerstadiation	*0*		1	14	
	I		5	93	
	IIa		5	126	
	IIb		2	27	
	IIIa		1	16	
	IIIb		5	83	
	IVa		4	41	
	IVb		1	17	
Cancerstage		*T 1*	5	29	
		2	3	83	
		3	12	237	
		4	7	109	
		N 0	16	278	
		1	7	118	
		2	5	76	
		M 0	21	401	
		1	7	73	
Cancergrading	*low*		5	95	
	high		17	237	
Lymphomononucleainfiltrate	*nonrefertable*		1	60	.481
	high		20	72	
	low		7	332	
Vascularinfiltrate	*nonrefertable*		3	26	.364
	present		23	334	
	absent		2	106	
Perinuclearinfiltrate	*nonrefertable*		8	31	.09
	present		17	255	
	absent		3	179	
MSI/MSS pattern	*lowprobability*		21	325	.493
	high probability		6	78	
Microsatellite analysis	MLH1	*notexpressed*	5	67	.213
		partialexpression	1	14	
		expressed	21	67	
	MSH2		0	11	.601
			1	17	
			26	11	
	MSH6		0	10	
			1	14	
			26	10	
	PMS2		6	69	
			0	13	
			21	69	
Oncogenesmutations	BRAF01	*mutated*	3	54	.069
		nonmutated	16	174	
	KRAS01		6	82	.327
			13	111	
	NRAS01		2	6	.031
			16	177	

The study recruited a total of 502 patients. Histological lymphomononuclear, perinuclear and vascular infiltration, as well as immunochemistry for MLH1, PMS2, MSH2, and MSH6 and hotspot mutational profiling of KRAS, BRAF, NRAS genes were defined.

1. Mutational analysis in young and old (>50 yrs) patients with colon cancer

To investigate the differences in the mutational rate between ages, the rate of the most frequently colon cancer mutated genes was compared in young and old (>50yrs) patients.

As shown in **Figure 3A**, the rate of NRAS mutation was significantly higher in young patients comparing to elderly patients (respectively 15% – 1%) (p=0.031).

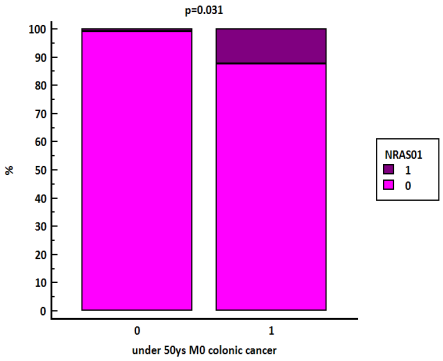


Figure 3A (for legend see Figure section)

KRAS gene analysis did not show any significantly different mutational rate between old and young patients, as shown in **Figure 3B** (30% young – 45% old pts) (p=0.327).

On the contrary, BRAF rate of mutation was found to be significantly lower in young than in old patients, with a rate mutation respectively of 5% in young and 30% in elderly subjects (p=0.069) (see **Figure 3C**).

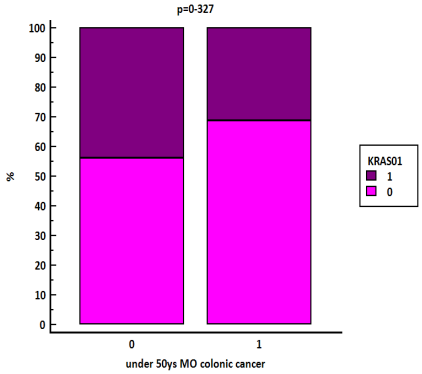


Figure 3B

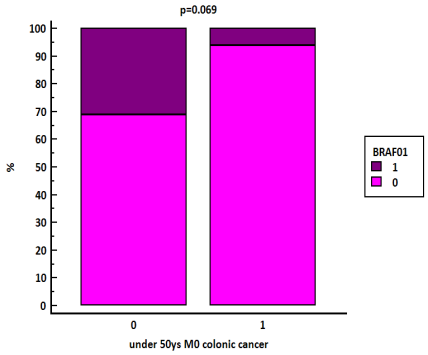


Figure 3C

2. *Differences in Microsatellite stability and instability of young vs. old patients (>50 yrs) with colon cancer*

To assess the differences in the microsatellite stability and instability rate, the mutations of the most frequently involved microsatellites were compared both in young and old (>50 yrs) patients with colon cancer.

For what concerns MSI and MSS cancer tumoral forms, the results showed a similar expressional pattern in microsatellite involvement between the two groups. A slightly infrequence in high probability microsatellite involvement was detected in young patients (15%young – >20% old pts) (see **Figure 4A**, for legend see Figure section). The two groups did not show any statistical difference (p=0.493).

The most frequently involved microsatellites were examined, showing a total absence of MSH2 loss of expression in young patients comparing to elderly ones (4% of MSH2 loss of expression detected). Also, a small amount of partial expression pattern of the same microsatellite was found in both groups in similar proportion (<5% of the total amount) (see **Figure 4B**). Statistical differences have not been detected between the two groups (p=0.601).

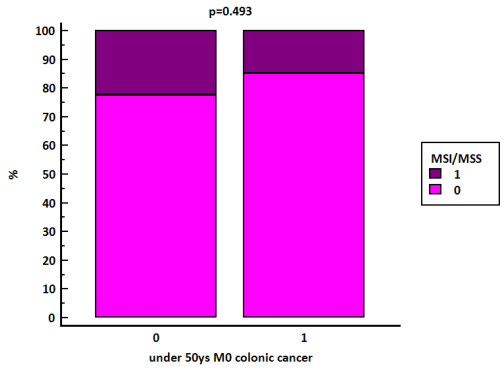


Figure 4A (MSI/MSS)

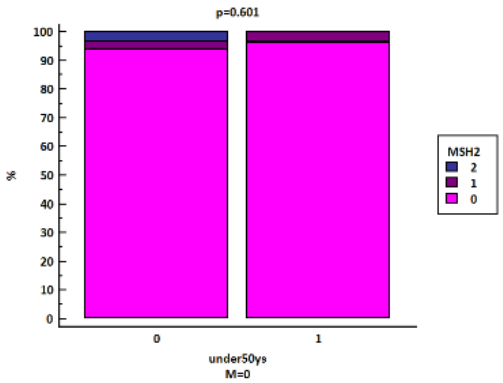


Figure 4B (MSH2)

Also, MLH1 expression pattern was analysed, showing a prevalence of its total absence in elderly patients comparing to young ones (respectively 20% in old and 11% in young), while partial microsatellite expression instead was shown prevalent in young patients (9% young vs. 2% old) (see **Figure 4C**). No statistical difference was seen (p=0.213).

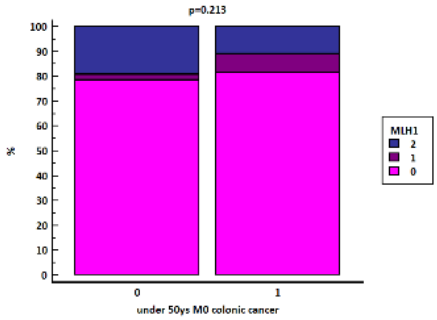


Figure 4C (MLH1)

3. *Differences in the histological tumor lymphomononuclear, vascular and perineural infiltrate between young and old patients (>50yrs) with colon cancer*

To investigate the histological tumor infiltrate comparing young and old patients (>50yrs) with colon cancer, lymphomononuclear, vascular and perineural involvement were considered.

Lymphomononuclear infiltrate was found to be low (purple column section, **Figure 5A**) in similar proportions both in young and old patients (respectively 69% and 70%), while the presence of a high infiltrate (blue column section, Figure 5A) and its absence (pink column section, Figure 5A) showed a complementary pattern result respectively in young and old patients, with a cumulative lower rate of absent infiltrate in young patients and a correspondent higher presence of high infiltrate in the same patients (respectively 9% none – 25% high infiltrate). On the contrary, old subjects were found to have a collective lower amount of high lymphomononuclear infiltrate than young ones, but a higher cumulative absence of the same parameter (respectively 12% none – 17% high infiltrate) (see **Figure 5A**). The two groups did not show any statistical difference (p=0.481).

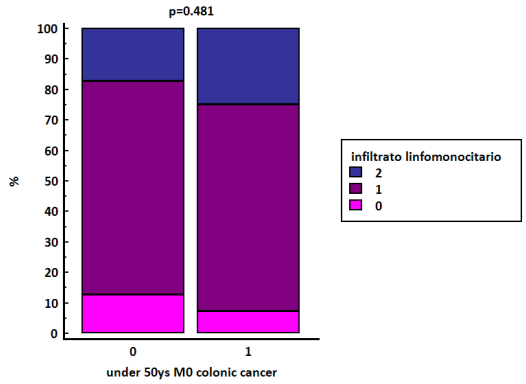
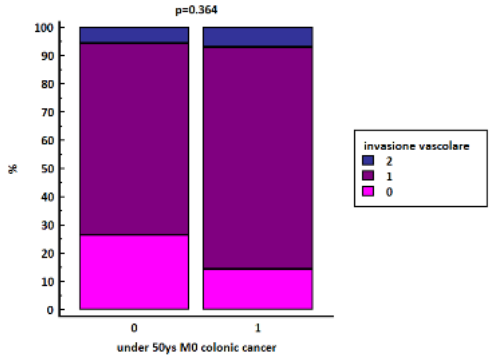


Figure 5A (for legend see Figure section)

Vascular invasion showed instead a different pattern of expression in its presence and absence between the two categories of patients, with a detectable prevalence of the invasion in young patients (about 80% young – 70% old patients) (purple section column, **Figure 5B**). Absence of infiltrate was more represented in old subjects instead (see Figure 5B, pink column). Respectively 5% of cases in both subjects could not be evaluated (blue section column, Figure 5B). No statistical difference between the two groups was detected (p=0.364).



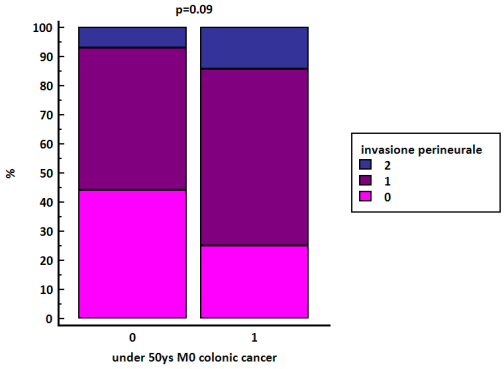


Figure 5B (for legend see Figure section) *Figure 5C*

The comparison of perineural invasion between young and old patients showed instead a similar proportion of the infiltrate presence (61% young – 40% old pts), while a different pattern was found for its absence, being almost doubled in old patients compared to young ones (respectively 44% and 25%) (see **Figure 5C**, purple column). Low statistical difference was found between the two groups (p=0.09).

4. *Differences in gene mutational profile in PanCancer Atlas (Colorectal Adenocarcinoma, TCGA)*

In order to have a wider perspective about mutational gene profile in colon cancer, tumor microenvironment was deeper investigated. The rate of the most frequently mutated genes and the expression levels of immunity involved genes were examined within the TCGA panCancer Atlas (2018) Colon Cancer series [21]. In particular, early (≤50 years) and elderly (>50 years) patient's age at the diagnosis were compared for the highest average frequency genes found mutated in colon cancer cells. Furthermore, the most significant p-values in gene alterations were considered.

As for gene mutation general profile, not great differences were found (no significant statistical difference detected). In particular, APC and TP53 resulted as the most frequently mutated genes between young and old patients, while young subjects especially were found to have a slightly higher mutational profile of TP53, TTN, KRAS, MUC16, CSMD3, SYNE1 genes (see **Figure 6A**). In particular, as for immune genes involved, MUC16 (that displays a role in lubricating barrier against particles and infectious agents at mucosal surfaces) [26] and CSMD1 (a potential suppressor of squamous cell carcinomas) [26] were found altered, respectively with a higher mutational rate within young patients and old ones (>50yrs) (see Figure 6A).

The most significant p-values in more commonly gene mutations were also analysed, not revealing significant statistic differences in young and elderly-onset colon cancer (see **Figure 6B**).

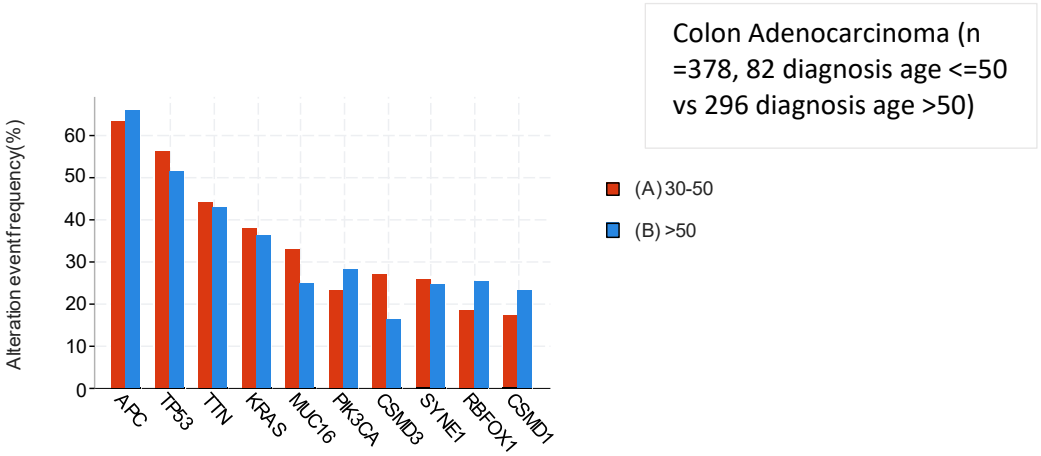


Figure 6A (Highest average frequency genes mutation)

Gene	(A) 30-50	(B) >50	p-Value	Enriched in
BTN2A2	7 (8.54%)	2 (0.60%)	2.490e-4	(A) 30-50

Table 2

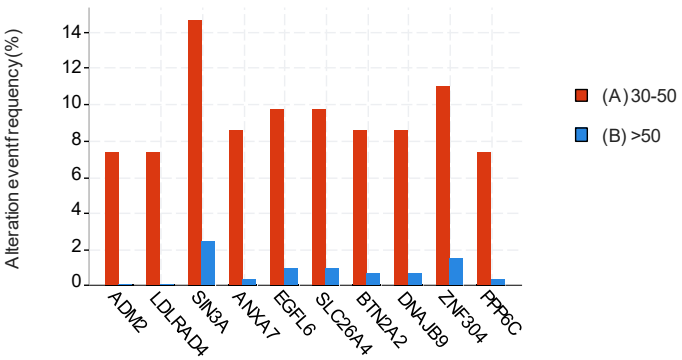


Figure 6B (*Most significant p-values in gene mutation*)

Expressly, as for the immune genes that showed the most significant statistical differences, BTN2A2 – known for its immune role of proliferation inhibitor of both CD4 and CD8 T-cells activated by anti-CD3 antibodies, regulator of T-cell metabolism and IL2 and IFNG secretion [26] – was found to have a strongly higher mutated pattern in young people (see **Table3**).

GENES	***IMMUNE***	***POSSIBLE IMMUNE INVOLVMENT***	***NOT IMMUNE***	***DESCRIPTION***
ADM2			.....	*This gene encodes a member of the calcitonin gene-related peptide (CGRP), a calcitonin family of hormones that play a role in the regulation of cardiovascular homeostasis, prolactin release, anti- diuresis, anti-natriuresis, and regulation of food and water intake.*
LDLRAD4		.....		*This gene functions as a negative regulator of TGF-beta signaling and thereby probably plays a role in cell proliferation, differentiation, apoptosis, motility, extracellular matrix production and immunosuppression. It is a signal transduction inhibitor.*
SIN3A		.....		*The protein encoded by this gene is a transcriptional repressor of MYC oncogenic activities. It has a role in regulating cell cycle progression.*
ANXA7			.....	*Annexin VII is a member of the annexin family of calcium-dependent phospholipid binding proteins (voltage-sensitive calcium channel activity, ion selectivity and membrane fusion), found in brain, heart and skeletal muscle.*
EGFL6			.....	*This gene encodes a member of the epidermal growth factor (EGF) repeat superfamily. It is often involved in the regulation of cell cycle, proliferation, and developmental processes. Expressed during early development, also detected in lung and meningioma tumors.*
SCL264			.....	*The SLC26A4 gene encodes an anion transporter known as pendrin and is the gene mutant in Pendred syndrome and enlarged vestibular aqueduct.*
BTN2A2	.....			***Inhibits the proliferation of CD4 and CD8 T-cells activated by anti-CD3 antibodies, T-cell metabolism and IL2 and IFNG secretion.***
DNAJB9		.....		*This gene is a member of the J protein family, regulating ATPase activity of heat shock proteins. The encoded protein is localized to the endoplasmic reticulum, is induced by endoplasmic reticulum stress and plays a role in protecting stressed cells from apoptosis.*
ZNF304		.....		*This gene encodes a member of the Krueppel C2H2-type zinc-finger family of proteins, funcioning as a transcriptional repressor (stimulates promoter hypermethylation and transcriptional silencing of target genes). Expression of this gene is upregulated in colorectal, ovarian and breast cancer. Also, this gene may promote cancer cell survival, growth and invasion.*
PPP6C			.....	*This gene encodes the catalytic subunit of protein phosphatase, a component of a signaling pathway regulating cell cycle progression.*

Table 3 (Immune genes involvment. Source: The Human Protein Atlas (2003)*)*

5. Differences in mRNA expression of most common lymphocyte markers in PanCancer Atlas (Colorectal Adenocarcinoma, TCGA)

In order to further investigate the immune role in tumor microenvironment infiltrate, an analysis on mRNA expression of most common lymphocyte markers was displayed using TCGA Colon Cancer Cases database (panCnacer Atlas – 2018) [21].

Among lymphocyte markers analysed, only GZMB, HLA-B, HLA-C and IL17B showed a significantly different expression between early and late-onset colon cancer cases.

As for the mRNA expression of lymphocyte GZMB marker – cytotoxic T-cells and NK-cells protease, activator of both caspase-independent pyroptosis and caspase-mediated apoptosis [26] – the group among 30-50 years age at the diagnosis showed a consistent reduced amount of expression (see **Figure 7A**) (p 0.0443).

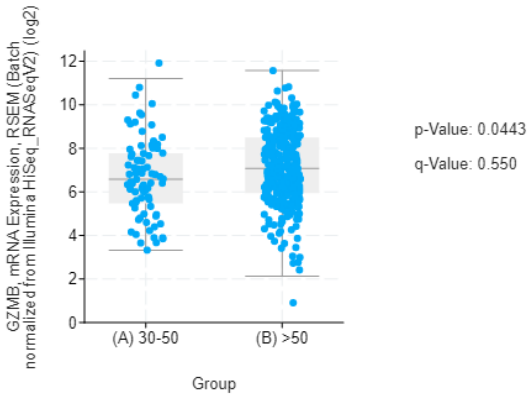


Figure 7A (GZMB)

Also, HLA-B (Major histocompatibility complex, Class I, B) and HLA-C (Major histocompatibility complex, Class I, C) – both guiding antigen-specific T cell immune response plus contributing to NK cells activity (HLA-C) [26] – showed a similarly reduced amount of expression in young groups (respectively **Figure 7B** and **7C**). Significant difference was found between young and elderly patients expression pattern of the two genes (respectively p 0.077 – p 0.0727 respectively).

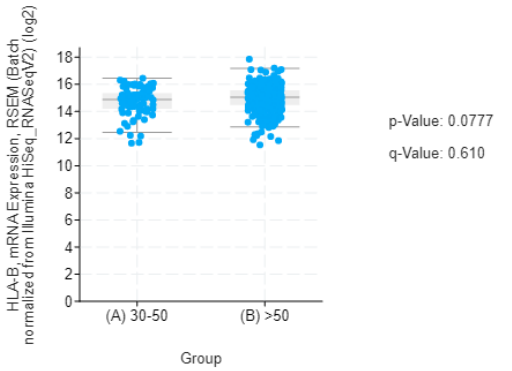


Figure 7B (HLA-B)

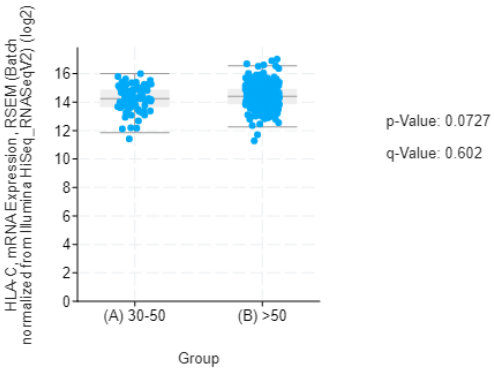


Figure 7C (HLA-C)

For about mRNA expression of IL17B instead – T cell-derived cytokine, involved in TNF-alpha and IL-1-beta releasing) [26] – the group among 30-50 years age at the diagnosis showed a significantly higher amount of expression compared to old population (see **Figure 7D**) (p 0.0174).

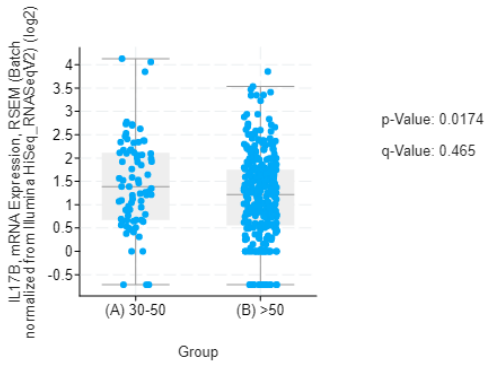


Figure 7D (IL17B)

Conclusions

Colorectal cancer during the last few decades has registered an increasing worrying growth, especially among young subjects (<50 years), in concurrence with the increasing of mortality in the same patients [4]. Although in early onset colon cancer genetic predisposition is shown to have a central role, most cases still remain sporadic [16]. Even so, the full spectrum of germline and somatic sequence variations implicated are unknown [16].

Changing in diet habits involving a major consume of red meat as well as well refined grains and processed sugar has been addressed as one of the main modifiable risk factors in early-onset CCR development [4][2][16]. Also, smoking, sedentary behaviour, overweight and obese body habitus are related to an increase of young CCR, together with antibiotic use, especially in early prenatal period and adolescence [2][4]. Latest theories present Exposome as a possible aetiology factor of sporadic young-onset CCR. It is described as the result of a combination of interactions between germline genetics and environment (GxE), and it is supposed to play a role in gut microbiota composition and consequent affection in sporadic CCR developing [2][16].

This project was designed to enlighten molecular and pathological features of young-onset colon cancer in order to have a deeper comprehension of its genesis. In order to achieve these goals, genetic and microsatellite profiles were analysed, together with immune infiltrate, mutated gene expression rate and lymphocytes markers density.

As for the oncogenes results, NRAS mutation was mainly detected in young-onset cancer profile rather than in old-one, while BRAF showed a significant mutation prevalence in old subjects. On the other side, KRAS mutations did not show any significant difference. These observations were also corroborated by other systematic analysis. In particular, Zubaidah at al [27] highlighted positive expression of BRAF mutation in young-onset colorectal cancer, as well as NRAS mutation (E132K), addressed as a qualified biomarker of young-onset CCR by Timothy et al [28]. Also, Ounissi et al suggested in their study NRAS and KRAS mutation profile as a crucial orientating therapy element [29].

An analysis of the most frequently mutated genes rate was displayed, showing in particular a mutation of MUC16, CSMD1 and BTN2A2 immune genes in young

subjects. For what concerns MUC16, it is known to be a cell surface-associated mucin, implicated to be upregulated in a large repertoire of malignancies, even though its function in the CRC pathogenesis still remains unknown [30]. Zhining et al in their study found that an overexpression of MUC16 in CRC patients was positively correlated with poor prognosis of patients [30]. Regarding CSMD1, a candidate tumor suppressor gene, it was found by Zhang et al as associated with overall survival [31]. Finally, as for BTN2A2, it is known to be a part of Btn-like (BTNL) molecules cluster, which have a role in T lymphocyte responses control and are genetically associated with inflammatory disorders and cancer [32]. In their study, Fernández et al found altered levels of BTN2A2 both in intestinal polyps of mice and colon inflammation [32].

Microsatellites and tumoral infiltrate showed similar patterns in the two groups, presenting the young-onset form a lower rate of high probability microsatellite involvement (especially lacking total absence of MSH2 expression). Recently, Gryfe et al [33] in their metanalyses found high-frequency microsatellite instability in colorectal cancer as an independently predictive factor of positive outcome and as linked to a reduce likelihood of metastases, addressing microsatellite analyses as a predictive marker of survival advantage [34][35][36].

Also, a predominant absence of the tumoral infiltrate pattern in all its three forms (lymphomononuclear, vascular and perineural) was observed in young patients. In addition, mRNA expression of most expressed lymphocyte markers was analysed, showing a general trend of reduced amount in young subjects (30-50 years at the diagnosis) for GZMB, HLA-B and HLA-C, leaving IL17B the only marker found higher in the same group of patients. Several studies demonstrated that antitumor immunity is associated with favourable prognosis in CRC patients [35]. Prizment et al. in their study demonstrated that higher tumor infiltration with CTL and GZMB cells is associated with improved all-cause and cancer-specific survival of CRC patients [38]. Furthermore, GzmB was found to provide cytotoxic activity against cancer cells, being a positive prognostic marker in human CRC [39][40]. Similarly, HLA-B and HLA-C high expression level was found as an independent predictor of favourable overall survival in colon and rectal cancer [41]. Interleukin-17 also, being a proinflammatory cytokine, has mostly been considered as promoter in CRC progression, promoting cancer inflammation and preventing tumoral cells from immune surveillance [42]. Recently however, in

their study Amicarella et al. found intraepithelial localisation of CRC-infiltration IL-17+ cells associated with improved survival [43].

Overall, a reduced GZMB, HLA-B and HLA-C together with an increased IL17B T cell infiltration in young subjects tumoral mucosa was found, suggesting a weak immune mucosa response as responsible for cancer neoantigens, opening the way to an early onset colon cancer.

The main limitations of our study are related to retrospective an observational design, population heterogeneity, absence of clinical data and outcomes association, and exclusion of metastatic cancer forms. Also, a lack of differentiation between lymphocytes and monocytes was displayed, as the retrospective series lymphomonocyte infiltration was obtained through conventional histology only (ie, hematoxylin and eosin staining). Moreover, as a small absolute number of NRAS and BRAF positive patients were detected, caution in interpreting results must be displayed.

As colorectal cancer among young subject is registering a worldwide increase, no actual explanation has been found so far, being this issue the central role of this study.

In conclusion, in our study mutational gene profile, immune infiltrate and microsatellite instability expression in early-onset colon cancer were found to have an overall similar trend of elderly-onset cases. However, deepening the study and focusing on immune gene alteration and lymphocyte infiltrate presence, tumoral mucosa of young-onset colon cancer was found to have an overall reduced immune expression of cytotoxic and regulatory lymphocytes, suggesting part of early onset colon cancer might be due to a weak immunesurveillance. Although further studies are needed, these findings, enlightening the crucial aspects of colon cancer genesis, lead the way to a better comprehension to this always more problematic reality.

Figures and Tables

Figure 1. *Study design and patients' selection for AOPD (Azienda Ospedale Università di Padova).*

Study diagram*. Flow chart showing the inclusion and exclusion of samples in the study.* ***Abbreviations****: AOPD, Azienda Ospedaliera di Padova; CC, colon cancer.*

Figure 2. *Study design and patients' selection for TCGA, PanCancer Atlas - Colorectal Adenocarcinoma database (2018).*

Study diagram*. Flow chart showing the inclusion and exclusion of samples in the study.* ***Abbreviations****: TCGA, The Cancer Genome Atlas.*

Figure 3 (A-B-C)*. Mutational analysis in young and old (>50 yrs) patients with colon cancer.*

Figure 3A *(NRAS01 mutation analyses): Column legend: 0: patients over 50 years; 1: patients under 50 yrs (≤50 years). Colour legend: 0: non mutated; 1: mutated.*

Figure 3B *(KRAS01 mutation analyses)****:*** *(See Figure 2A description).*

Figure 3C *(BRAF01 mutation analyses)****:*** *(See Figure 2A description).*

Figure 4 (A-B-C)*. Differences in Microsatellite stability and instability of young vs. old patients (>50 yrs) with colon cancer.*

Figure 4A *(MSI/MSS analyses)****:*** *Column legend: 0: patients over 50 years; 1: patients under 50 yrs (≤50 years). Colour legend: 0: low probability; 1: high probability.*

Figure 4B *(MLH1 analyses)****:*** *Column legend: (see above). Colour legend: 0: total expression; 1; partial expression; 2: absent expression*

Figure 4C *(MSH2 analyses)****:*** *(See Figure 3B description).*

Figure 5 (A-B-C)*. Differences in the histological tumor lymphomononuclear, vascular and perineural infiltrate between young and old patients (>50yrs) with colon cancer.*

Figure 5A *(Lymphomononuclear infiltrate analyses)****:*** *Column legend: 0: patients over 50 years; 1: patients under 50 yrs (≤50 years). Colour legend: 0: not present: 1: low infiltrate; 2: high infiltrate.*

Figure 5B *(Perivascular infiltrate analyses)****:*** *Column legend: (see above). Colour legend: 0: not present; 1: present; 2: not valuable.*

Figure 5C *(Perinuclear infiltrate analyses)****:*** *(See Figure 4B description).*

Figure 6 (A-B)*. Differences in gene mutational profile in PanCancer Atlas (Colorectal Adenocarcinoma, TCGA)*

Figure 6A *(Highest average frequency genes mutation*)

Figure 6B (*Most significant p-values in gene mutation*)

Table 2. *(BTN2A2 gene datas and p-value)*

Table 3. *(Immune genes involvment)*

Figure 7 (A-B-C-D) *Differences in gene mutational profile in PanCancer Atlas (Colorectal Adenocarcinoma, TCGA)*

Figure 7A *(GZMB mRNA expression of T lymphocyte)*

Figure 7B *(HLA-B mRNA expression of T lymphocyte)*

Figure 7C *(HLA-C mRNA expression of T lymphocyte)*

Figure 7D *(IL17B mRNA expression of T lymphocyte)*

References

1. Evelien Dekker, Pieter J Tanis, Jasper L A Vleugels, Pashtoon M Kasi, Michael B Wallace. Colorectal cancer. *The Lancet*, 2019; 394: 1467–80.

2. Naohiko Akimoto, Tomotaka Ugai, Rong Zhong, Tsuyoshi Hamada, Kenji Fujiyoshi, Marios Giannakis, Kana Wu, Yin Cao, Kimmie Ng and Shuji Ogino. Rising incidence of early- onset colorectal cancer — a call to action. *Nature*, 2021; 18: 230-243.

3. Firas Baidoun, Kholoud Elshiwy, Yasmine Elkeraie, Zahi Merjaneh, George Khoudari, Muhammad Talal Sarmini, Mohamed Gad, Muneer Al-Husseini, Anas Saad. Colorectal Cancer Epidemiology: Recent Trends and Impact on Outcomes. *Bentham Science Publishers*, 2021;22(9):998-1009.

4. Frank A. Sinicrope. Increasing Incidence of Early-Onset Colorectal Cancer. *New England Journal of Medicine*, April 2022;386:1547-58.

5. Siegel RL, Torre LA, Soerjomataram I, et al. Global patterns and trends in colorectal cancer incidence in young adults. *Gut* 2019; 68: 2179-85.

6. Vuik FE, Nieuwenburg SA, Bardou M, et al. Increasing incidence of colorectal cancer in young adults in Europe over the last 25 years. *Gut* 2019; 68: 1820-6.

7. Siegel RL, Miller KD, Goding Sauer A, et al. Colorectal cancer statistics, 2020. CA *Cancer J Clin* 2020; 70: 145-64.

8. Araghi M, Soerjomataram I, Bardot A, et al. Changes in colorectal cancer incidence in seven high-income countries: a population-based study. *Lancet Gastroenterol Hepatol* 2019; 4: 511-8.

9. Hidetsugu Yamagishi, Hajime Kuroda, Yasuo Imai and Hideyuki Hiraishi. Molecular pathogenesis of sporadic colorectal cancers. *Chinese Journal of Cancer*, 2016;35:4.

10. Franck Pagès, M.D., Ph.D., Anne Berger, M.D., Ph.D., Matthieu Camus, M.Sc.,Fatima Sanchez-Cabo, Ph.D., Anne Costes, B.S., Robert Molidor, Ph.D,Bernhard Mlecnik, M.Sc., Amos Kirilovsky, M.Sc., Malin Nilsson, B.S.,Diane Damotte, M.D., Ph.D., Tchao Meatchi, M.D., Patrick Bruneval, M.D., Ph.D., Paul-Henri Cugnenc, M.D., Ph.D., Zlatko Trajanoski, Ph.D., Wolf-Herman Fridman, M.D., Ph.D., and Jérôme Galon, Ph.D. Effector Memory T Cells, Early Metastasis, and Survival in Colorectal Cancer. *The New England Journal of Medicine*, 2005; 353:2654-66.

11. Veroushka Ballester, Shahrooz Rashtak, Lisa Boardman. Clinical and molecular features of young-onset colorectal cancer. *World Journal of Gastroenteroly,* 2016; 22(5): 1736-1744.

12. Matthew B. Yurgelun, MD; Serena Masciari, MD; Victoria A. Joshi, PhD; Rowena C. Mercado,MD, MPH; Noralane M. Lindor, MD; Steven

Gallinger, MD; John L. Hopper, PhD; Mark A. Jenkins, PhD; Daniel D. Buchanan, PhD; Polly A. Newcomb, PhD, MPH; John D. Potter,MD, PhD; RobertW. Haile, PhD; Raju Kucherlapati, PhD; Sapna Syngal, MD, MPH; for the Colon Cancer Family Registry. Germline TP53 Mutations in Patients With Early-Onset Colorectal Cancer in the Colon Cancer Family Registry. *JAMA Oncology,* May 2015;1(2):214-221.

13. Liang JT, Huang KC, Cheng AL, Jeng YM, Wu MS, Wang SM. Clinicopathological and molecular biological features of colorectalcancer in patients less than 40 years of age. *British Journal of Surgery,* 2003; 90:205-214.

14. Antelo M, Balaguer F, Shia J, Shen Y, Hur K, Moreira L, Cuatrecasas M, Bujanda L, Giraldez MD, Takahashi M, Cabanne A, Barugel ME, Arnold M, Roca EL, Andreu M, Castellvi-Bel S, Llor X, Jover R, Castells A, Boland CR, Goel A. A high degree of LINE-1 hypomethylation is a unique feature of early-onset colorectal cancer. *PLoS One*, 2012; 7: e45357.

15. Des C. Winter, Centre for Colorectal Disease, St Vincent's University Hospital, Dublin, Ireland. REACCT Collaborative Members. Microsatellite instability in young patients with rectal cancer: molecular findings and treatment response. *British Journal of Sociology*, 2022; 109, 251–255.

16. Alexandra M. Zaborowski, MD, Centre for Colorectal Disease, St Vincent's University Hospital, Elm Park, Dublin 4, Ireland. REACCT Collaborative Authors. Characteristics of Early-Onset vs Late-Onset Colorectal Cancer. *JAMA Surgery*, 2021;156(9):865-874.

17. Wild CP. Complementing the genome with an "exposome": the outstanding challenge of environmental exposure measurement in molecular epidemiology. *Cancer Epidemiololgy Biomarkers & Prevention*, 2005;14(8):1847-1850.

18. 17. Wild CP. Complementing the genome with an "exposome": the outstanding challenge of environmental exposure measurement in molecular epidemiology. *Cancer Epidemiololgy Biomarkers & Prevention*, 2005;14(8):1847-1850.

19. Cancer Genome Atlas Research Network. MSK, the Dana Farber Cancer Institute, Princess Margaret Cancer Centre in Toronto, Children's Hospital of Philadelphia, The Hyve in the Netherlands, and Bilkent University in Ankara, Turkey. Colorectal Adenocarcinoma (*TCGA, PanCancer Atlas*). TCGA, Cell 2018.

20. Remo A, Fassan M, Lanza G. Immunohistochemical evaluation of mismatch repair proteins in colorectal carcinoma: *The AIFEG/GIPAD proposal. Pathologica*. 2016; 108:104e109.

21. Bocker T, Diermann J, Friedl W, et al. Microsatellite instability analysis: A multicenter study for reliability and quality control. *Cancer Research,* 1997; 57: 4739e4743.

22. Simbolo M, Gottardi M, Corbo V, et al. DNA qualification workflow for next generation sequencing of histopathological samples. *PLoS One*. 2013;8: e62692.

23. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: *R Foundation for Statistical Computing*; 2016.

24. Li Ding, Matthew H. Bailey, Eduard Porta-Pardo, Vesteinn Thorsson, Antonio Colaprico, Denis Bertrand, David L. Gibbs, Amila Weerasinghe, Kuan-lin Huang, Collin Tokheim, Isidro Corte´ s-Ciriano, Reyka Jayasinghe, Feng Chen, Lihua Yu, Sam Sun, Catharina Olsen, Jaegil Kim. Perspective on Oncogenic Processes at the End of the Beginning of Cancer Genomics. *Cell* April 5, 2018; 173, 305–320.

25. The Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* 487, 330–337 (2012). Supplementary Information Research (www.nature.com/nature), 11252: 1-55.

26. The Human Protein Atlas (2003). Uhlén et al (2015). Tissue-based map of the human proteome. *Science.* Human Protein Atlas available from *www.proteinatlas.org*/*www.cancer.gov/tcga*.

27. Zubaidah Saizul, Ab Hamid Siti-Azrin, Andee Dzulkarnaen Zakari, Aniza Hassan, Wan Faiziah Wan Abdul Rahman, Nur Asyilla Che Jalil. BRAF V600E and Mismatch Repair Proteins Expression in Sporadic Young-onset Colorectal Cancer in Kelantan, Malaysia. *Oman Medicine Journal*, 2006 July ;38(7):787-93.

28. Ryan Timothy D Yu, Reynaldo L Garcia. NRAS mutant E132K identified in young-onset sporadic colorectal cancer and the canonical mutants G12D and Q61K affect distinct oncogenic phenotypes. *Scientific Reports,* 2020 Jul 3;10(1):11028.

29. Donia Ounissi, Marwa Weslati, Rahma Boughriba, Meriam Hazgui, Saadia Bouraoui. Clinicopathological characteristics and mutational profile of KRAS and NRAS in Tunisian patients with sporadic colorectal cancer. *Turkish Journal of Medical Sciences*, 2021 Feb 26;51(1):148-158.

30. Zhining Liu, Yimei Gu, Xianghua Li, Lianbang Zhou, Xiaohu Cheng, Heng Jiang, Yang Huang, Yingfeng Zhang 2, Tongtong Xu, Wanshui Yang, Qiang Huang. Mucin 16 Promotes Colorectal Cancer Development and Progression Through Activation of Janus Kinase 2. *Digestive Diseases and Sciences*, 2021.

31. Rui Zhang, Chun Song. Loss of CSMD1 or 2 may contribute to the poor prognosis of colorectal cancer patients. *International Journal of General Medicine*, 2021; 14 8689–8710.

32. Cristina Lebrero-Fernández, Ulf Alexander Wenzel, Paulina Akeus, Ying Wang, Hans Strid, Magnus Simrén, Bengt Gustavsson, Lars G Börjesson, Susanna L Cardell, Lena Öhman, Marianne Quiding-Järbrink, Anna Bas-Forsberg. Altered expression of Butyrophilin (BTN) and BTN-like (BTNL) genes in intestinal inflammation and colon cancer. *Immunity Inflammatory Disease*, 2016 Apr 1; 4(2):191-200.

33. Robert Gryfe, m.d., Hyeja Kim, M.sc., Eugene t.k. Hsieh, M.D., Melyssa D. aronson, M.SC., Eric J. Holowaty, M.D., Shelley B. Bull, Ph.d., Mark Redston, M.D., and Steven Gallinger, M.D. Tumor microsatellite instability and clinical outcome in young patients with colorectal cancer. *The New England Journal of Medicine,* 2000; 342 (2): 69-77.

34. Boland CR, Thibodeau SN, Hamilton SR, et al. A National Cancer Institute workshop on microsatellite instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. *Cancer Res* 1998; 58:5248-57.

35. Ko JM, Cheung MH, Kwan MW, et al. Genomic instability and alterations in Apc, Mcc and Dcc in Hong Kong patients with colorectal carcinoma. *International Journal of Cancer* 1999; 84:404-9.

36. Salahshor S, Kressner U, Fischer H, et al. Microsatellite instability in sporadic colorectal cancer is not an independent prognostic factor. *British Journal of Cancer* 1999; 81:190-3.

37. Jass, J. R. Lymphocytic infiltration and survival in rectal cancer. *Journal of Clinical Pathology,* 1986; 39, 585–589.

38. Anna E. Prizment, Robert A. Vierkant, Thomas C. Smyrk, Lori S. Tillmans, Heather H. Nelson, Charles F. Lynch, Thomas Peng, Stephen N. Thibodeau, Timothy R. Church, James R. Cerhan, Kristin E. Anderson, and Paul J. Limburg. Cytotoxic T-cells and granzyme B associated with improved colorectal cancer survival in a prospective cohort of older women. *Cancer Epidemiology Biomarkers and Prev*ention, 2017 April; 26(4): 622–631.

39. Llipsy Santiago, Marta Castro, Rebeca Sanz-Pamplona, Marcela Garzo´ n, Ariel Ramirez-Labrada, Elena Tapia, Vı´ctor Moreno, Elena Layunta, Gabriel Gil-Go´ mez, Marta Garrido, Rau´ l Pen˜ a, Pilar M. Lanuza, Laura Comas, Paula Jaime-Sanchez, Iratxe Uranga-Murillo, Rosa del Campo, Pablo Pelegrı´n, Eric Camerer, Luis Martı´nez-Lostao, Guillermo Mun˜oz, Jose´ A. Uranga, Anabel Alcalde, Eva M. Galvez, Angel Ferrandez, Phillip I. Bird, Sunil Metkar, Maykel A. Arias and Julian Pardo. Extracellular GranzymeA Promotes Colorectal Cancer Development by Enhancing Gut Inflammation. *Cell Reports* 32, 107847, July 7, 2020.

40. Galon, J., Page` s, F., Marincola, F.M., Angell, H.K., Thurin, M., Lugli, A., Zlobec, I., Berger, A., Bifulco, C., Botti, G., et al. Cancer classification using the Immunoscore: a worldwide task force. *Journal of Translational Medicine,* 2012; 10, 205.

41. Theodoros Michelakos, Filippos Kontos, Tomohiro Kurokawa, Lei Cai, Ananthan Sadagopan, Danielle Krijgsman, Wilko Weichert, Lindy G Durrant, Peter J K Kuppen, Cristina R Ferrone, Soldano Ferrone. Differential role of HLA-A and HLA-B, C expression levels as prognostic markers in colon and rectal cancer. *The American Journal of Surgery Pathology*, 2007 Jan; 31(1):76-84.

42. Dang Wu, Pin Wu, Qi Huang, Yang Liu, Jun Ye, and Jian Huang. Interleukin-17: A Promoter in Colorectal Cancer Progression. *Hindawi Publishing Corporation, Clinical and Developmental Immunology*, 2013; 1-7.

43. Amicarella, M G Muraro, C Hirt, E Cremonesi, E Padovan, V Mele, V Governa, J Han, X Huber, R A Droeser, M Zuber, M Adamina, M Bolli, R Rosso, A Lugli, I Zlobec, L Terracciano, L Tornillo, P Zajac, S Eppenberger-Castori, F Trapani, D Oertli, G Iezzi. Dual role of tumour-infiltrating T helper 17 cells in human colorectal cancer. *Gut*, 2017; 66:692–704.