Article

Transcriptome Analyses Identify Deregulated *MYC* in Early Onset Colorectal Cancer

Olivia M. Marx [1,2], **Marc M. Mankarious** [2] , **Melanie A. Eshelman** [3], **Wei Ding** [2], **Walter A. Koltun** [2] **and Gregory S. Yochum** [1,2,*]

1 Department of Biochemistry & Molecular Biology, Pennsylvania State University College of Medicine, Hershey, PA 17033, USA
2 Department of Surgery, Division of Colon & Rectal Surgery, Pennsylvania State University College of Medicine, Hershey, PA 17033, USA
3 Department of Pediatrics, Division of Hematology & Oncology, Pennsylvania State University College of Medicine, Hershey, PA 17033, USA
* Correspondence: gsy3@psu.edu

Abstract: Despite a global decrease in colorectal cancer (CRC) incidence, the prevalence of early-onset colorectal cancer (EOCRC), or those occurring in individuals before the age of 50, has steadily increased over the past several decades. When compared to later onset colorectal cancer (LOCRC) in individuals over 50, our understanding of the genetic and molecular underpinnings of EOCRCs is limited. Here, we conducted transcriptomic analyses of patient-matched normal colonic segments and tumors to identify gene expression programs involved in carcinogenesis. Amongst differentially expressed genes, we found increased expression of the *c-MYC* proto-oncogene (*MYC*) and its downstream targets in tumor samples. We identified tumors with high and low differential *MYC* expression and found patients with high-*MYC* tumors were older and overweight or obese. We also detected elevated expression of the *PVT1* long-non-coding RNA (lncRNA) in most tumors and found gains in copy number for both *MYC* and *PVT1* gene loci in 35% of tumors evaluated. Our transcriptome analyses indicate that EOCRC can be sub-classified into groups based on differential *MYC* expression and suggest that deregulated *MYC* contributes to CRCs that develop in younger patients.

Keywords: early-onset colorectal cancer; RNA-sequencing; transcriptome; MYC; PVT1

1. Introduction

Cancers of the colon and rectum are estimated to rank third in terms of newly diagnosed cancers and are expected to be the third-leading cause of cancer-related deaths [1]. While the overall colorectal cancer (CRC) incidence has been in decline since the early 1990s, there has been a steady increase in the incidence of early-onset colorectal cancer (EOCRC) [2,3]. EOCRCs are primarily detected as a result of abdominal pain or rectal bleeding, as opposed to later-onset colorectal cancers (LOCRCs), which are commonly found through surveillance colonoscopy [4]. Though the majority of EOCRCs are diagnosed between the ages of 40–49 [5], the American Cancer Society only recently (2018) began recommending that CRC screenings begin at age 45 instead of 50 [6]. As such, at the time of diagnosis, EOCRC patients frequently present with later-stage tumors and often have metastatic disease [7].

While genetic predisposition in terms of family history and hereditary conditions are associated with approximately a third of EOCRC cases [2], our understanding of the genetic and molecular pathways that drive carcinogenesis in these patients is far from complete. Compared with LOCRC, EOCRCs are most frequently microsatellite stable, near diploid, and often have an increased mutational burden [2]. Furthermore, several studies indicate that there are differences in the frequency and underlying mutations in oncogenic drivers in EOCRCs versus LOCRCs [2,8–12]. Despite these distinctions, current treatments

for EOCRC and LOCRC remain largely the same [13]. To identify biomarkers and new treatment strategies for EOCRC and to better understand why there continues to be an increase in incidence worldwide, more work is needed to define the molecular and genetic underpinnings of EOCRC.

Expression of the *c-MYC (MYC)* proto-oncogene is deregulated in over 50% of all cancers [14]. MYC is a basic-helix-loop-helix transcription factor that binds to MAX to control the expression of numerous genes involved in cell cycle progression, metabolism, cell survival, and ribosome biogenesis [15–18]. Overexpression of MYC has been shown to contribute to early carcinogenesis in cancers, including prostate [19] and gastric [20]. Furthermore, studies suggest that MYC overexpression can promote field cancerization, which is thought to be a cancer-initiating event in which tissue regions are altered to promote increased proliferation [21]. Mechanisms that contribute to *MYC* deregulation in cancers include retroviral insertions, chromosomal translocations, locus amplification, and hyper-activation of oncogenic signal transduction pathways [22,23]. In addition, recent studies implicated the long-non-coding RNA (lncRNA) *PVT1* as a regulator of oncogenic MYC [24–26]. *PVT1* is located approximately 65 kb downstream from *MYC* at the 8q24 chromosomal locus, and it is frequently co-amplified in those tumors displaying gains in *MYC* copy number [25]. Moreover, *PVT1* binds and stabilizes MYC which in turn activates *PVT1* gene expression in a positive feedback loop to promote carcinogenesis [24,25]. Though it is well documented in the literature that MYC is a driver of colorectal cancer [23,27], its potential role in EOCRCs is less clear.

Large publicly available cancer datasets such as The Cancer Genome Atlas provide valuable resources for genomic and transcriptomic sequences of cancers. However, for many cancers, these resources lack patient-matched, non-cancerous samples [28]. The inclusion of patient-matched normal tissues is necessary to control for demographic or tissue-specific factors that could contribute to altered gene expression patterns in tumors [28,29]. Here, we use genome-wide RNA-sequencing (RNA-seq) on surgically resected tumors and adjacent colonic segments in a cohort of patients with EOCRC. We uncover evidence that the expression of *MYC*, and its downstream targets, are upregulated in tumors and find that tumors can be sub-classified based on differential *MYC* expression. We also detected elevated expression of *PVT1* in most tumor samples and gains in the *MYC/PVT1* genomic locus in a third of the samples. We found that those patients who had high-*MYC* expressing tumors at the time of surgery were older in our cohort and were obese or overweight. For the first time, these results implicate *MYC* as an oncogene in EOCRC.

2. Materials and Methods

2.1. Specimen Collection

This study was approved by the Pennsylvania State University College of Medicine Institutional Review Board (IRB Protocol No. PRAMSHY98-057). Prior to surgery, patients gave informed consent to have surgically resected intestinal tissues and blood to be collected and stored in the Inflammatory Bowel and Colorectal Disease Biobank (IBCRD) that is maintained within the Department of Surgery at the Pennsylvania State University College of Medicine. Patients in the study were under the age of 50 years old at the time of surgery, which was performed at our institution between 2015–2021. A flowchart of patient selection is included in Supplemental Figure S1. Blood was drawn prior to surgery, and resected tissues were collected from the gross pathology lab and transported to our lab for processing. Full-thickness sections of rectum, cecum, sigmoid colon, ascending colon, and descending colon were collected, and aliquots were frozen in liquid nitrogen, fixed in formalin, and stored in RNAlater (Thermo Fisher, Waltham, MA, USA) preservative or embedded in paraffin for immunofluorescence analysis. Microsatellite stability was assessed through immunohistochemical staining of tumor samples for nuclear expression of four mismatch repair (MMR) proteins MLH1, MSH2, MSH6, and PMS2. Tumors were characterized as microsatellite unstable if defects in expression of MMR proteins were

detected. Optional genetic testing using a microarray platform was performed for patients with microsatellite instability to confirm and diagnose Lynch syndrome.

2.2. Indirect Immunofluorescence

Formalin-fixed tissue samples were embedded in paraffin and cut into 5 μm thick slices and placed onto VWR superfrost plus slides. Slides were dewaxed with xylene (two 5 min washes) and rehydrated with ethanol and water-graded washes (100% ethanol for 3 min, twice; 95% ethanol for 3 min, twice; 70% ethanol, once; 50% ethanol, once; water 5 min). Antigen retrieval was performed in 10 mM citrate buffer (pH 6) in a steamer for 20 min. Cooled slides were washed with PBS and blocked for 1 h in 5% donkey serum. Slides were incubated with primary antibodies against c-MYC (1:200 dilution, Cell Signaling Technology antibody #D84C12) in 5% donkey serum overnight. Slides were then washed in PBS. Secondary antibodies (Alexa Fluor 568 donkey anti-rabbit, Invitrogen, Waltham, MA, USA, #A10042) and DAPI (ThermoFisher #62248) were incubated on slides for 1.5 h. Slides were washed with PBS and mounted with mounting buffer (Invitrogen #P36982). Confocal fluorescent images were obtained at 63× magnification. Cells from each slide section were scored from 0–3 based on nuclear MYC staining intensity, and each cell value was averaged for a final score. This analysis was performed in triplicate for each sample.

2.3. RNA-Sequencing

Tumors and adjacent full-thickness colonic tissues (n = 21 pairs) were collected in an Eppendorf microcentrifuge tube containing RNAlater and homogenized with a pestle. Chloroform was added, and the samples were centrifuged at 12,000× g for 15 min. The aqueous layer was subjected to further purification using an RNAEasy Mini Kit (Qiagen, Germantown, MD, USA), and RNA integrity values were assessed using a Bioanalyzer (Agilent, Santa Clara, CA, USA). Samples were further selected for transcripts that were poly-adenylated, and cDNA libraries were sequenced on a MiSeq2500 (Illumina, San Diego, CA, USA) sequencer as described previously [30–32]. Paired-end reads were generated at a read length of 100 nucleotides. Read quality was assessed with the program FastQC (v11.8, Babraham Bioinformatics, Cambridge, UK), all Phred scores were above 30, and GC bias was not detected in the samples. An average of 32 million reads per sample were aligned to the human genome (build hg38) using STAR version 2.7.3 (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA) with default parameters [33]. Aligned reads were counted using HTSeq [34]. Samples contained an average of 88% uniquely mapped reads. Quality assessment of the RNA-seq libraries is presented in Supplementary Figure S2.

2.4. Hierarchical Clustering and Differential Gene Expression Analysis

The R package DESeq2, version 1.32.0 (Boston, MA, USA) [35], was used to normalize gene expression values and to identify differential expression of genes between all tumor and all colonic control samples. Differentially expressed genes were hierarchically clustered and visualized with gene-scaling for each row using the pheatmap function in R. Gene set enrichment analysis (GSEA) for cancer hallmarks was performed with 1000 phenotype permutations on rlog normalized counts from DESeq2 output on all tumors compared to all controls [36]. Gene ontology (GO) categories were assessed in R using package EnsDb.Hsapiens.v86 to convert ENSEMBL gene IDs to ENTREZ ID or gene symbol. The clusterprofiler package was used to identify ontology groups, and Enrichplot was used to visualize ontology results in R [37,38]. In order to calculate the log2 fold-change of expression in each tumor versus the patient-matched control, the DESeq2 function rlog was first used to determine the log2 normalized expression values, and then gene expression values from the adjacent colonic tissue were subtracted from values obtained from the corresponding tumor.

2.5. Cell Type Deconvolution

In order to estimate the proportions of cell types in each sample, cellular deconvolution was performed as previously described [31] using xCell [39]. Results were plotted in R using ggplot2 [40]. We calculated significance between cell proportion scores using a paired Wilcoxon test for tumor versus adjacent normal samples, and we used an unpaired Wilcoxon test to compare high and low-*MYC* expressing tumors.

2.6. Analysis of Transcriptomes in The Cancer Genome Atlas (TCGA)

Transcriptome data were retrieved from the TCGA COAD (colorectal adenocarcinoma) study (TCGA Research Network. Available online: https://www.cancer.gov/tcga accessed on 1 October 2021) using TCGAbiolinks package in R [41]. A total of 521 HTSeq-Counts files were filtered to isolate 474 primary tumor samples. Age at diagnosis under 18,250 days (approximately 50 years), was considered EOCRC (*n* = 55), and age above 18,250 days was considered LOCRC (*n* = 419). Count data were normalized using DESeq2 [35] and EnsDb.Hsapiens.v86 was used to convert ENSEMBL gene IDs to gene symbols.

2.7. Reverse Transcription and Quantitative PCR

cDNA was synthesized from 400 ng of purified RNA from each patient sample using the Verso cDNA synthesis kit (Thermo Fisher, Waltham, MA, USA) with random hexamer primers according to the manufacturer's instructions. cDNAs were diluted to 10 ng/µL and subjected to PCR amplification with SYBR green chemistry using reactions and parameters previously described [30]. The following oligonucleotide sequences were used in the reactions: *MYC*; forward, AACTTGACCCTCTTTGGCAGCA and reverse, GCAAACCTCCTCACAGCCCAC; *β-ACTIN*; forward, GAGCATCCCCCAAAGTTCACAATG and reverse TGGCTTTTAGGATGGCAAGGGACT. *MYC* values were normalized to *β-ACTIN*, and fold change in expression was calculated using $2^{-\Delta Ct}$. The resulting qPCR values were compared with the normalized *MYC* expression values from the RNA-Seq analysis, and Pearson correlation was used to assess agreement between the approaches.

2.8. Copy Number Variation Analysis

Genomic DNA was isolated from 5–25 mg of flash-frozen tissues and tumors using the QIAamp DNA minikit (Qiagen, Germantown, TN, USA) according to the manufacturer's instructions. Quantitative PCR was performed using 20 ng of genomic DNA as the template in 20 µL reactions using the QuantStudio3 Real Time PCR system (Applied Biosystems, Waltham, MA, USA). The *MYC* genomic locus was evaluated using the Taqman probe Hs02602824_cn (Thermo Fisher), and the *PVT1* genomic locus was evaluated using Taqman probe Hs06232749_cn. The *RNASEP* gene (Thermo Fisher, 4403326) served as an internal reference. Reactions were performed in quadruplicate, and non-template controls were included. Results were analyzed with CopyCaller (Applied Biosystems, version 2.1).

2.9. Statistical Analyses

Statistical analyses were performed using R software v4.1.1. Two-tailed statistical tests were determined to be significant at $p < 0.05$. Significance of clinical correlates were determined using Fisher's Exact test. Pearson's correlation was used to determine the relationship between gene expressions. Spearman's correlation was used to determine the relationship between gene expression and copy number variation. Paired Wilcoxon tests were used to compare gene expression in patient-matched normal colonic segments and tumors. All methods were carried out in accordance with relevant guidelines and regulations.

3. Results

3.1. Patient Cohort

To define genes with EOCRC-associated expression, we first identified a cohort of patients in our inflammatory bowel and colorectal disease biobank who were under the age of 50 and for which we had archived flash-frozen tumors and adjacent uninvolved full-thickness colonic segments. For simplicity, we will refer to the adjacent colonic samples as

control segments. We found 21 such patients whose demographics are listed in Table 1. In summary, our cohort is predominantly (90%) Caucasian, 57% male, and 76% overweight or obese, and these patients did not receive chemotherapy prior to surgery. Five patients (24%) had additional intestinal diseases, with four patients (19%) diagnosed with diverticulitis and one (5%) with inflammatory bowel disease. Tumors were resected throughout the colon and rectum, although no tumors were available from the transverse colon. An immediate family history of colorectal cancer was reported in nine patients (43%). Four patients (19%) exhibited defects in mismatch repair proteins, two of whom were diagnosed with Lynch syndrome after further analyses. All cancers were adenocarcinomas, with the majority being stage III or IV at the time of resection.

Table 1. Patient Demographics.

Parameter	N	(%)
Sex		
Male	12	(57)
Female	9	(43)
Race		
White	19	(90)
Black	1	(5)
Asian	1	(5)
BMI		
Underweight	0	(0)
Normal	5	(24)
Overweight	4	(19)
Obese	12	(57)
History of Diabetes	1	
History of smoking [a]	8	(38)
Family history of colorectal cancer [b]	9	(43)
Location		
Right Colon	6	(29)
Left Colon	1	(5)
Sigmoid	6	(29)
Rectosigmoid	2	(10)
Rectum	6	(29)
Stage		
I	6	(29)
II	3	(14)
III	10	(48)
IV	2	(10)
Adenocarcinoma	21	(100)
Signet/Mucinous histology	4	(19)
Lymphovascular invasion	7	(33)
Peri-neural invasion	6	(29)

[a] Positive smoking status includes both current and former smokers at the time of surgery; [b] Positive family history for colon or rectal cancer in first degree family member.

3.2. Tumors and Adjacent Colonic Segments Display Distinct Transcriptomic Profiles

We isolated poly-adenylated RNAs from tumor and control samples and conducted RNA-seq to identify gene signatures associated with EOCRC. Using hierarchical clustering and principal component analysis (PCA) of normalized gene expressions of the top 4000 most variable genes in all samples, we found that with respect to their transcriptomes, tumors and controls form distinct clusters, with few exceptions (Figure 1A,B). Patient numbers 1 and 4 were diagnosed with Lynch syndrome, and patients 9 and 12 were found to have MMR defects but were not diagnosed with Lynch syndrome. In addition, tumor transcriptomes did not cluster based on the region of the colon from which they were resected (Figure 1A). We next used DESeq2 to identify differentially expressed genes in tumors versus controls and found that 1521 were upregulated (log2FoldChange > 1.5,

adjusted $p < 0.05$) and 1542 were downregulated (log2FoldChange < -1.5, p-adjusted < 0.05) (Figure 1C). Gene ontology (GO) analysis on genes whose expression was significantly up-regulated (log2FoldChange > 1.5, p-adjusted < 0.05) in tumors identified processes involved with nuclear division, chromosomal maintenance, and signaling receptor activity, amongst others (Figure 1D, Supplementary Table S1). The EOCRC GO signature is characteristic of highly proliferative cells. Cell type deconvolution analysis was performed to predict cellular populations of normal and tumor cells, confirming that the largest proportion of cells analyzed were either epithelium or smooth muscle cells. This analysis also identified differences in immune cell populations in tumors compared with normal samples (Supplementary Figure S3).
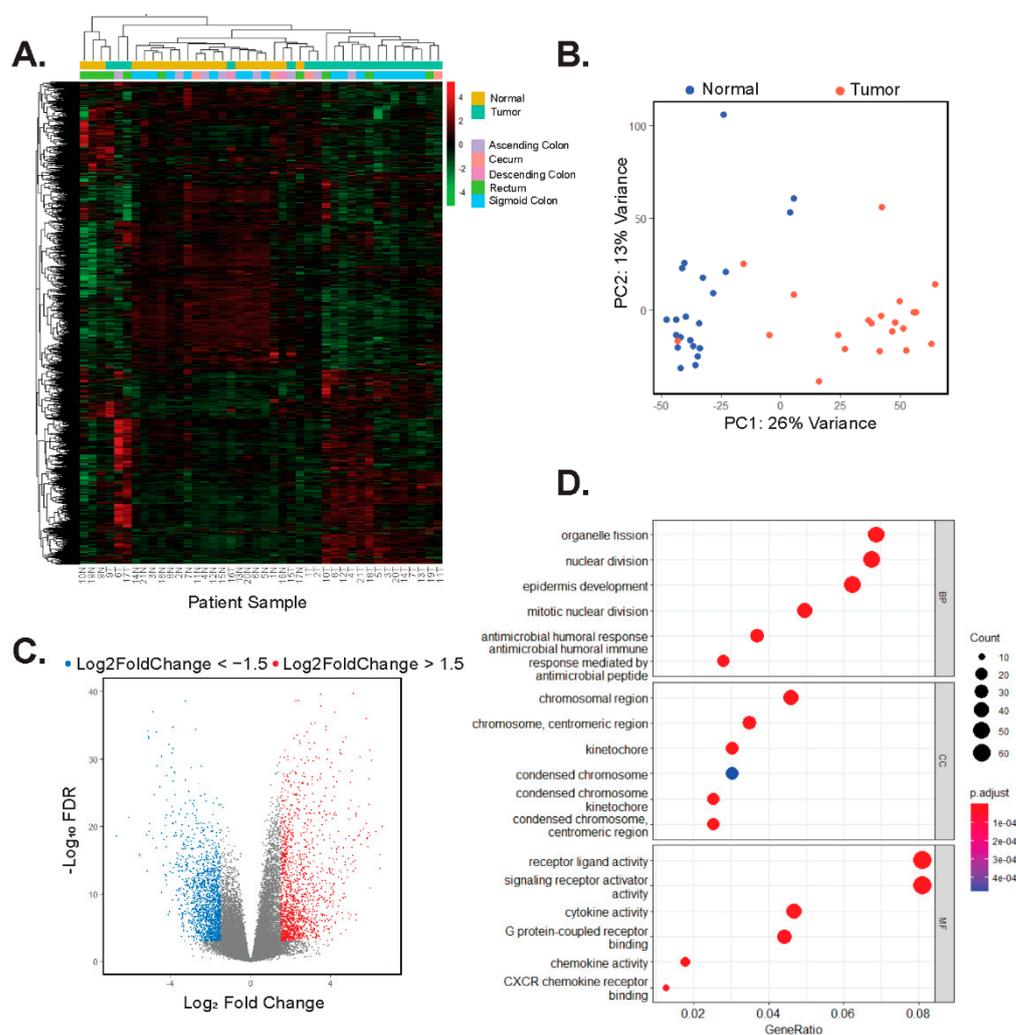

Figure 1. Tumors and adjacent colonic segments display distinct transcriptomic profiles. (**A**) Heatmap and hierarchical clustering of rlog normalized gene expression of the top 4000 most variable genes in tumors and adjacent control tissues. Patients are arbitrarily assigned 1–21 and N represents uninvolved tissues, whereas T represents tumor. (**B**) Principal component analysis (PCA) of the top 4000 most variable genes expressed in uninvolved (normal) and tumors. (**C**) Volcano plot of differentially expressed genes in tumors versus uninvolved control colonic tissues. Red points are 1521 genes whose expression is significantly upregulated (log2Foldchange > 1.5, p-adj < 0.05) and blue points are 1542 genes whose expression is significantly down-regulated (log2Foldchange < -1.5, p-adj < 0.05) in tumors versus controls. (**D**) GO analysis on the 1521 genes whose expression is upregulated in tumor samples. The top six results in each of the categories, biology processes (BP), cellular component (CC), and molecular function (MF) are shown in descending order.

3.3. MYC and Its Downstream Target Genes Are Differentially Expressed in EOCRC

We next applied gene set enrichment analysis (GSEA) to identify genes, or pathways, associated with hallmarks of cancer that were differentially expressed in the tumors versus controls [36,42]. Surprisingly, the expression of genes comprising the Wnt/β-catenin signaling hallmark set [42] was not significantly enriched in tumors as deregulation of this pathway has been previously described for both EOCRC and LOCRC [43,44] (Supplementary Figure S4A–C). Instead, we found significant enrichment (*p* < 0.05) of three hallmarks of cancer: unfolded protein response, DNA damage repair (Supplementary Figure S4D–G), and MYC targets (Figure 2A, Supplementary Table S2). We decided to further investigate MYC, given its established role as a driver of colorectal carcinogenesis [23,27]. We found that in addition to the enrichment of its downstream targets, the *MYC* oncogene itself was significantly enriched in tumors, and its differential expression clearly distinguished tumor samples from controls (Figure 2B). Through indirect immunofluorescence (IF) analysis, we found that expression of MYC is increased in tumors versus adjacent controls (Figure 2C,D). These findings suggest that deregulated expression of *MYC*, and its downstream targets, are involved in early-onset colorectal cancer pathogenesis.
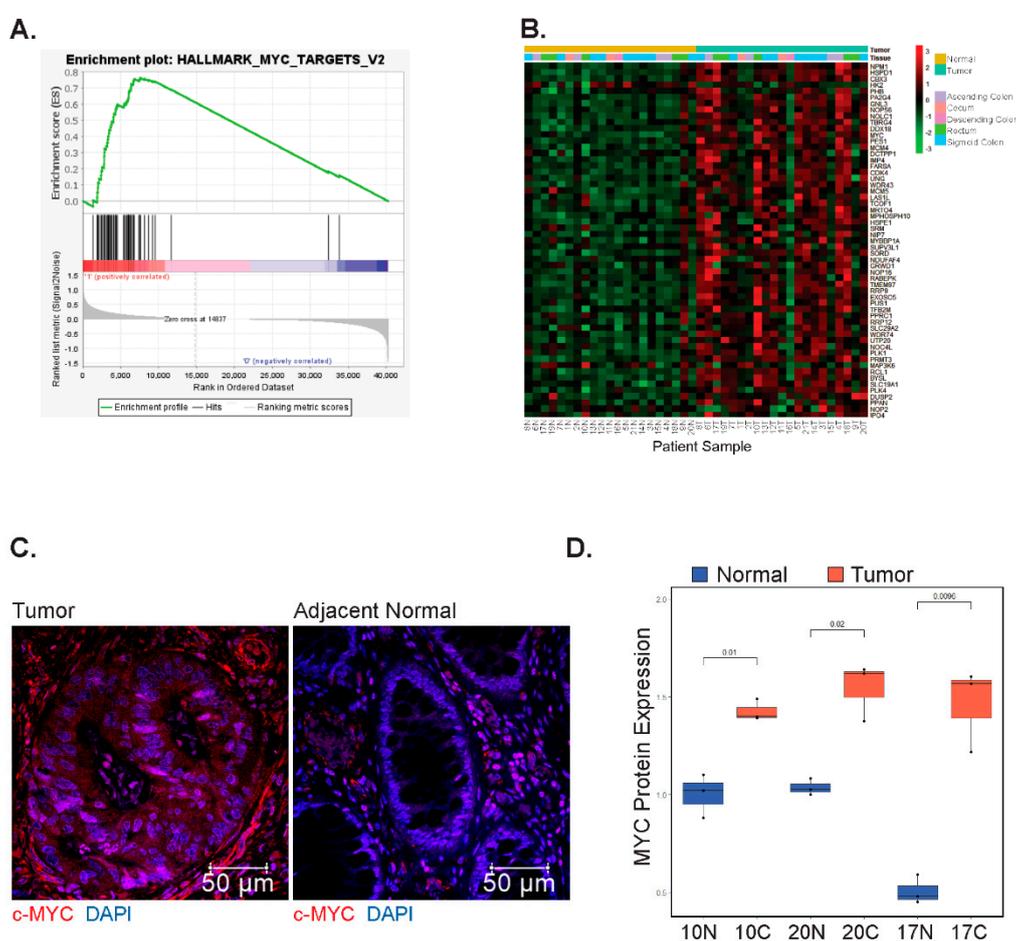

Figure 2. *MYC* and its downstream targets are differentially expressed in EOCRC. (**A**) Gene set enrichment analysis (GSEA) of genes comprising the hallmark MYC target set in tumors compared with patient-matched uninvolved colonic segments (Nominal *p*-value = 0.011, Normalized Enrichment Score: 1.52). (**B**) Heatmap of row-scaled gene expression of *MYC* and its downstream targets in tumors and adjacent colonic segments. Patients are arbitrarily assigned 1–21 and N represents uninvolved tissues, whereas T represents tumors. (**C**) Representative IF images of tumors and matched normal samples stained using a primary antibody against MYC. (**D**) Quantification of *MYC* positive nuclei from 3 fields of view per sample (N = 3 matched samples).

3.4. Differential MYC Target Gene Expression Clusters Patients into Two Groups

We assessed whether the expression of genes comprising the MYC target hallmark set could identify sub-populations of patients in our study. To address this possibility, we used DESeq2 to normalize the data and to quantify fold-change expression differences of MYC hallmark genes in each patient-matched tumor and control tissue. Upon hierarchical clustering of these values, we found two discrete sub-classes of patients. The first subclass, comprising thirteen patients (62%), displayed increased expression of *MYC* and its downstream targets (Figure 3A). We will refer to these patients and tumors from these patients as high-*MYC*. In the second subclass of eight patients (38%), expression of *MYC* and most of its downstream targets were either down-regulated or unchanged in patient-matched tumors versus control segments (Figure 3A). We will refer to these patients, and their corresponding tumors, as low-*MYC*. Notably, as we grouped patients based on differential expression of *MYC*, and its downstream targets, tumors in the low-*MYC* group may still exhibit comparable levels of *MYC* transcripts to those expressed in tumors of the higher *MYC* group. As an independent validation of *MYC* expression, we performed RT-qPCR on isolated tumor and control RNAs. Plotting of RT-qPCR results versus normalized RNA-seq read counts confirmed the identification of high and low tumor profiles with respect to *MYC* expression (Figure 3B). Transcriptome data downloaded from The Cancer Genome Atlas (TGCA) of tumors from 55 colorectal adenocarcinoma (COAD) patients under 50 years of age show similar clustering into groups based on *MYC* hallmark signature gene expression, whereas in the over 50 age group, more samples have an intermediate expression of *MYC* targets (Supplementary Figure S5).

When considering our patient population as a whole, *MYC* expression is increased in all but two patient tumors versus matched control segments (Figure 3C). We used principal component analysis (PCA) of the top 5000 most variably expressed genes in tumors and found that those comprising the high-*MYC* sub-type were largely segregated from those in the low-*MYC* group (Figure 3D). When comparing high and low *MYC* tumors, we found that *MYC* is significantly upregulated in high-*MYC* tumors compared with low-*MYC* tumors (Wilcoxon test, $p = 9.8 \times 10^{-6}$, Supplementary Figure S6A). To further differentiate high and low-*MYC* tumors, cell deconvolution and GO analysis were performed on significantly ($p < 0.05$) deregulated genes in the high-*MYC* tumors compared with the low-*MYC* tumors. Notably, immune response gene signatures were downregulated in high-*MYC* compared with low-*MYC* tumors (Supplementary Figure S6B–D). In addition, we found that the expression of direct *MYC* target genes, including those involved in the polyamine synthesis pathway, were upregulated in the high- versus low-*MYC* tumors ($p < 0.05$, Supplementary Figure S7). Importantly, when considering several clinical correlates, patients in our cohort with high-*MYC* tumors tended to be older and overweight or obese (Table 2).

3.5. MYC and PVT1 Copy Number Alterations Are Found in Some Patient Tumors

Using RNA-seq data alone, it is difficult to assess which signal transduction pathways may impinge upon *MYC* to induce its expression. Because we found no associations in our samples with deregulated Wnt/β-catenin signaling, we sought to investigate other mechanisms which may contribute to increased *MYC* in the tumor samples. The *MYC* gene locus is subjected to copy number gains in numerous cancers, including CRC [45]. We, therefore, purified genomic DNA from matched tumors and controls from 17 patients and assessed the status of the *MYC* gene locus using a specific Taqman probe in quantitative PCR reactions. A probe against the *RNASEP* gene was included as an internal reference. We found that tumors from six patients (35%) displayed *MYC* copy number gains, with *MYC* copy numbers of 2.5 or higher (Figure 4A). Importantly, each of these six was among the high-*MYC*-expressing tumor population in our study. Tumors with copy number gains had increased levels of lymphovascular invasion, and they presented in older patients within our cohort (Supplementary Table S3). Notably, several other tumors were found to have gained in the 2.2–2.5 range, while all normal samples were within the 1.8–2.2 range. Six of the twelve tumors tested in the high-*MYC* sub-type did not contain greater than

2.5 copies of *MYC,* whereas none of the low-*MYC* tumors exhibited changes in copy number relative to uninvolved tissue (Figure 4B).
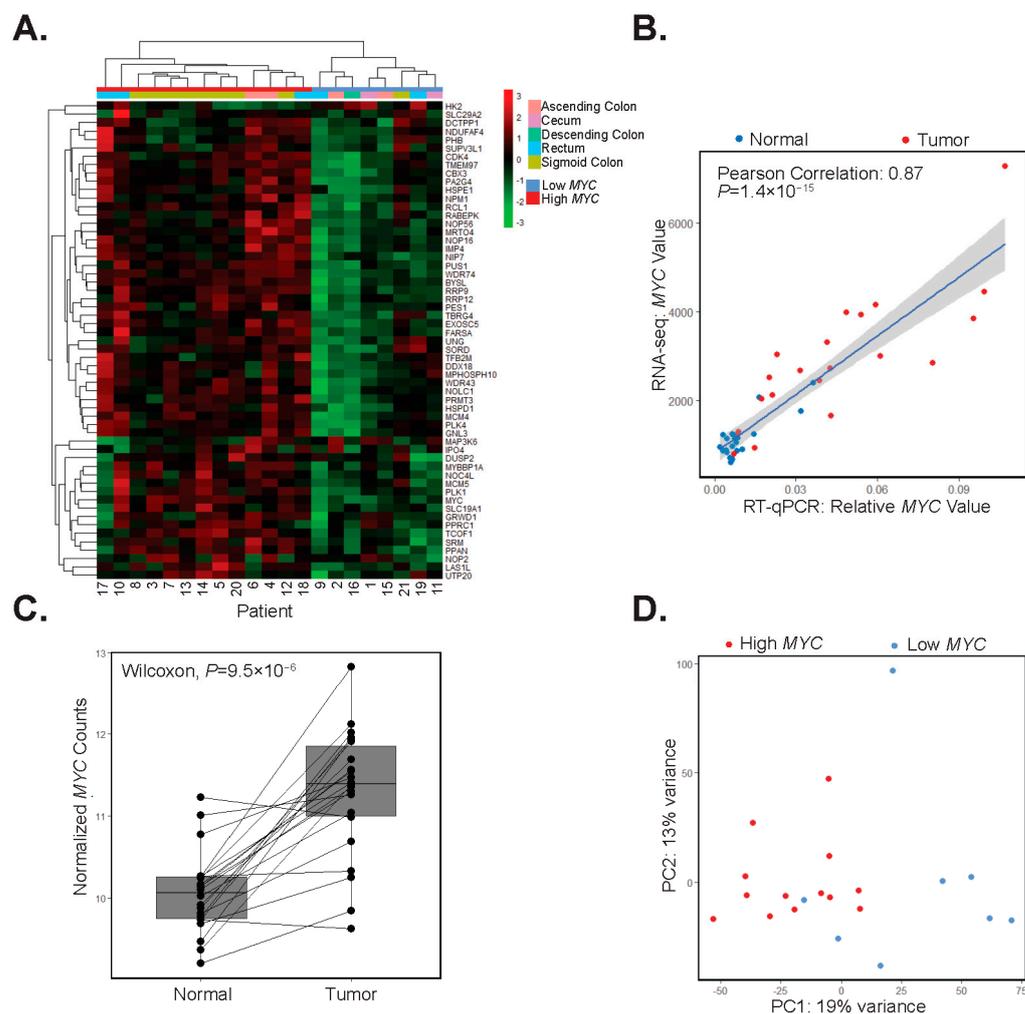

Figure 3. Differential MYC target gene expression clusters patients into two groups. (**A**) Hierarchical clustering of differential expression of *MYC* and its target genes in patient-matched tumors and adjacent colonic segments scaled by row. (**B**) RT-qPCR analysis of *MYC* and correlation with its expression in RNA-seq datasets in adjacent colonic segments (normal, blue points) and tumors (red points) (Pearson correlation 0.89, $p = 1.4 \times 10^{-15}$). (**C**) Boxplot of rlog normalized expression of *MYC* with lines connecting adjacent colonic segments (Normal) and tumor samples (paired Wilcoxon test, $p = 9.5 \times 10^{-6}$). (**D**) Principal component analysis of the top 5000 most variable genes in tumor samples with red points indicating high-*MYC* and blue points indicating low-*MYC* tumors.

To further investigate amplification in the 8q24 chromosomal region, we interrogated the *PVT1* locus in genomic DNA isolated from tumors displaying *MYC* copy number gains and their corresponding controls. We found that the *PVT1* locus was co-amplified in each of these tumor samples (Figure 4C). Together, these results suggest that *MYC* and *PVT1* copy number alterations account for increased *MYC* expression in half of the high-*MYC* tumors in our patient population.

Table 2. Patient demographics for low and high differential *MYC* expression.

Parameter	Low *MYC* Expression	High *MYC* Expression	*p*-Value
Patients, n	8	13	
Age at Diagnosis, median (IQR)	40 (39–44)	47 (46–48)	**0.043**
Sex			0.37
Male, n (%)	6 (75%)	6 (46%)	
Female, n (%)	2 (25%)	7 (54%)	
Race			0.13
White, n (%)	6 (75%)	13 (100%)	
Black, n (%)	1 (13%)	0 (0%)	
Asian, n (%)	1 (13%)	0 (0%)	
BMI			**0.047**
Normal or underweight, n (%)	4 (50%)	1 (8%)	
Overweight or Obese, n (%)	4 (50%)	12 (92%)	
History of Smoking [a], n (%)	4 (50%)	4 (31%)	0.65
Family History [b], n (%)	1 (13%)	8 (62%)	0.067
Located at Sigmoid or Rectum, n (%)	3 (38%)	11 (85%)	0.056
Overall Stage			0.53
I, n (%)	1 (13%)	5 (38%)	
II, n (%)	2 (25%)	1 (8%)	
III, n (%)	4 (50%)	6 (46%)	
IV, n (%)	1 (13%)	1 (8%)	
Signet/Mucinous component	4 (50%)	0 (0%)	**0.012**
Lymphovascular Invasion	3 (38%)	4 (31%)	1.00
Peri-neural Invasion	3 (38%)	3 (23%)	0.63

[a] Positive smoking status includes both current and former smokers at the time of surgery; [b] Positive family history for colon or rectal cancer in first degree family member. Bold values represent statistically significant differences between high-MYC and low-MYC groups (Fisher's exact test, $p < 0.05$).

3.6. Expression of PVT1 Correlates with MYC Expression

PVT1 has been shown to potentiate oncogenic MYC function in CRC [24–26]. Since we have shown that *MYC* and *PVT1* are co-amplified in some tumors, we next evaluated *PVT1* expression in our RNA-seq dataset. We found that *PVT1* expression was increased in most tumors (90%) relative to adjacent control tissue (Figure 5A). Moreover, for most samples, *PVT1* and *MYC* gene expressions correlated (Figure 5B). In addition, we noted a correlation between differential expression of both *PVT1* and *MYC* in tumors versus control tissues in patient-matched samples (Figure 5C). These results suggest that regardless of copy number status, *PVT1* may be a positive effector of oncogenic *MYC* in early-onset colorectal carcinogenesis.
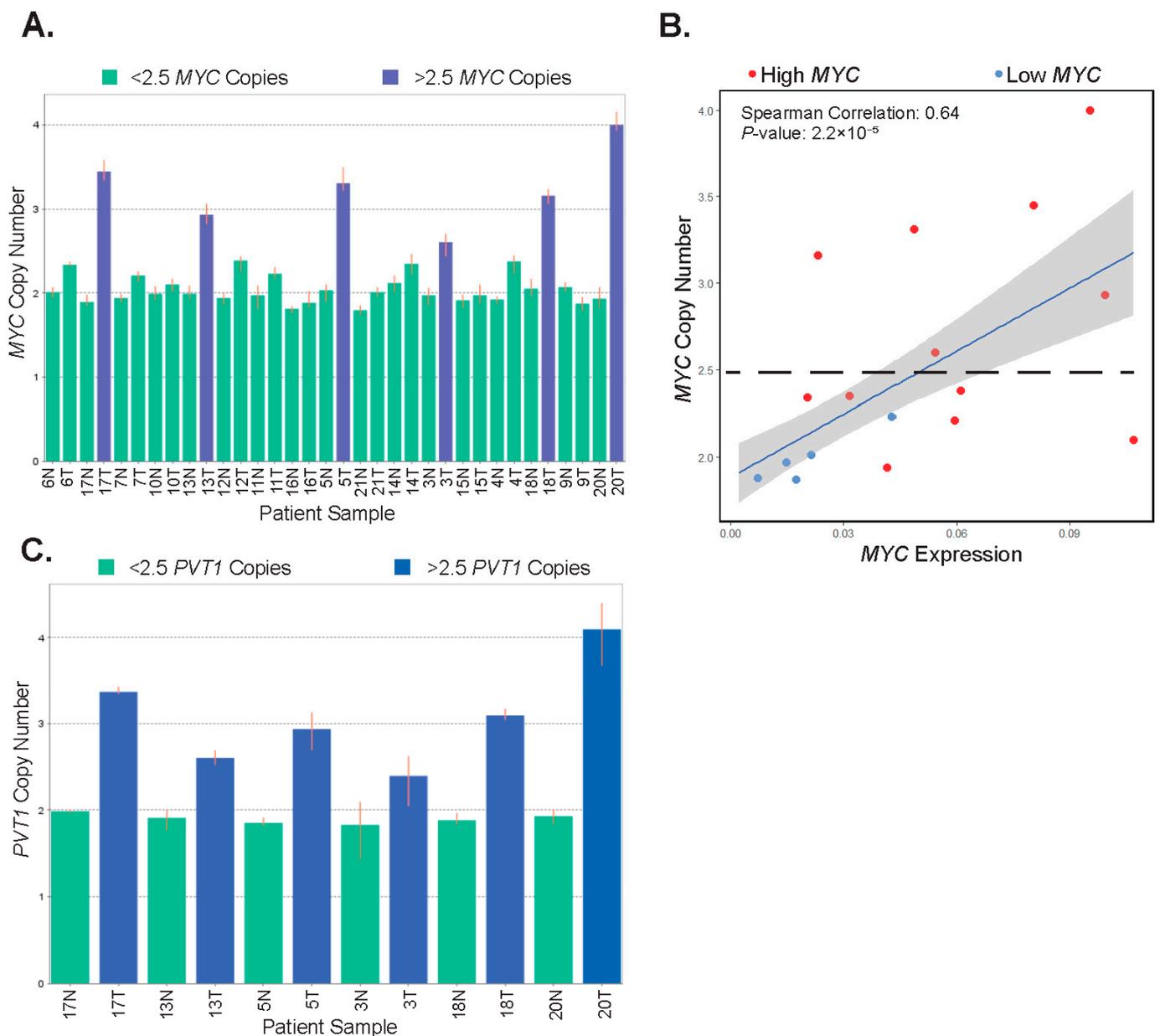
Figure 4. *MYC* and *PVT1* copy number alterations are found in some patient tumors. (**A**) Bar graph depicting the copy number of the *MYC* gene locus in patient-matched uninvolved colonic segments (normal, N) and tumors (T). Highlighted in blue are tumor samples with 2.5 or greater copy number. (**B**) Correlation of *MYC* copy number and *MYC* expression in tumors. Red points are tumors expressing high-*MYC* and blue points are tumors expressing low-*MYC* (Pearson correlation estimate 0.68, $p = 5.45 \times 10^{-6}$). (**C**) Bar graph depicting the copy number of the *PVT1* locus in the six patient-matched samples displaying gains in the *MYC* genomic locus in tumors (T). Uninvolved colonic tissues (normal, N) are shown as a reference. In (**A**,**C**), highlighted in blue are tumors samples with 2.5 or greater increases in copy number. Error bars represent the range of copy numbers detected in four technical replicates per sample.
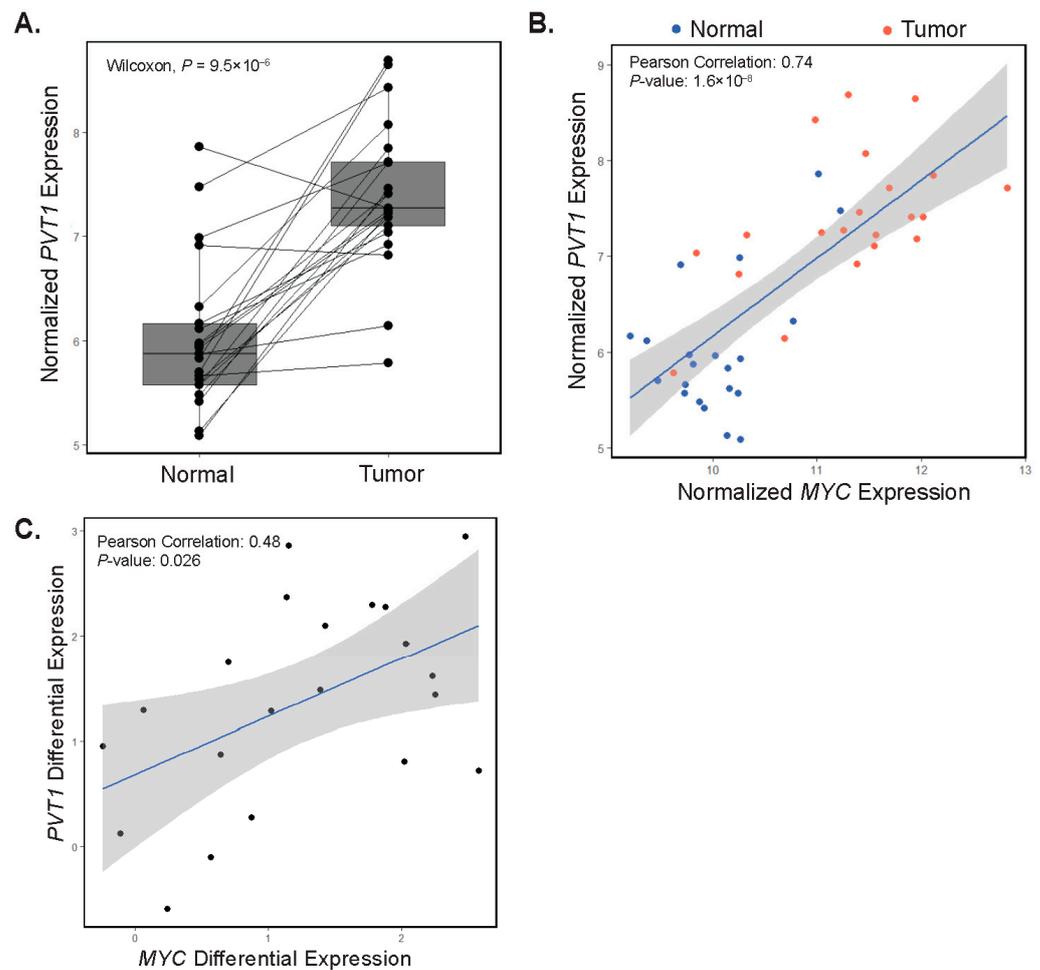
Figure 5. Expression of *PVT1* correlates with *MYC* expression. (**A**) rlog normalized expression plot of *PVT1* in patient-matched adjacent colonic tissues (Normal) and tumors (paired Wilcoxon test, $p = 4.8 \times 10^{-6}$). Each point represents a single patient with a line drawn between matched samples. (**B**) Correlation plot depicting normalized *MYC* versus normalized *PVT1* gene expression in adjacent colonic segments (Normal, blue points) and tumors (red points). (**C**) Correlation plot depicting *PVT1* and *MYC* differential expression in patient-matched adjacent colonic segments and tumors.

4. Discussion

In this study, we conducted RNA-seq analysis of patient-matched tumors and controls to identify molecular signatures associated with EOCRC. Hierarchical clustering of transcriptomes demonstrated that the vast majority of tumors clustered separately from control samples. That most tumors cluster separately from uninvolved tissues is in line with a previous study that reported transcriptomic profiles in organoid cultures derived from colonic mucosa and EOCRC [43]. Similarly, an earlier study reported that EOCRC gene expression profiles clustered separately from healthy controls [46]. Together, these results indicate that colonic gene expression profiles are significantly altered during the cellular transformation process to promote early-onset colorectal carcinogenesis.

Mutations in components of the Wnt/β-catenin signaling pathway are prevalent in CRC, and aberrant Wnt/β-catenin signaling activates *MYC* expression [47]. As in LOCRC, mutations in the APC tumor suppressor have been reported in studies of EOCRC; however, such mutations were found in approximately 70% of tumors surveyed, which is lower than the 80–90% frequency reported in LOCRC [43,48]. Yan et al. reported deregulated expression of some Wnt/β-catenin targets in most tumor-derived EOCRC organoids in their study, which supports Wnt/β-catenin as a driver of early-onset tumors [43]. In contrast to this finding, our GSEA of the Wnt/β-catenin hallmark target gene set failed to identify the dereg-

ulation of Wnt hallmark genes in the tumor set as a whole (Supplementary Figure S4A). One difference between our study and that of Yan et al. is that we conducted transcriptomics on flash-frozen tumors and full-thickness control segments, whereas theirs was conducted on organoids, which are derived from the colonic epithelium [43]. It is possible that expression profiles obtained from surrounding stromal or immune cell infiltrate in our samples mask a Wnt/β-catenin target gene signature in epithelial cells. Indeed, cell deconvolution analysis predicts that epithelial cells contribute only part of the gene signature, with smooth muscle, immune, and stromal cells, among other cell types, contributing as well (Supplementary Figure S3). While we did not assess the mutational status of *APC* or other Wnt-signaling pathway genes in our analysis, the lack of a deregulated Wnt/β-catenin target gene signature argues against this pathway as a significant driver of EOCRC in our surveyed population.

Our GSEA identified *MYC* and its downstream targets as the top hallmark of cancer pathway that distinguishes early-onset versus control transcriptomes. To our knowledge, this is the first report to implicate increased *MYC* expression in EOCRC. Though somatic mutations in the *MYC* gene have been reported in EOCRC, whether these impacted *MYC* expression levels was not evaluated [48,49]. Moreover, there are conflicting reports on whether somatic *MYC* mutations are more common in earlier or later onset CRC [48,49]. Hierarchical clustering of our samples based on the differential of MYC hallmarks separated our patients into two clusters, higher and lower *MYC* expression. Future work will determine whether the somatic mutations in *MYC* influence its expression levels in the tumors of our cohort.

We detected elevated expression of characterized MYC target genes in the high-*MYC* versus low-*MYC* tumors and found increased MYC protein expression via immunofluorescent staining of three patient's tumors and adjacent normal tissues, arguing that MYC protein expression is also increased and that it is directly deregulating its downstream targets. Notably, we identified increases in MYC targets associated with polyamine synthesis, including ornithine decarboxylase 1 (*ODC1*), spermidine synthase (*SRM*), and adenosylmethione decarboxylase 1 (*AMD1*) [50] (Supplementary Figure S7). Previous studies show aberrant MYC signaling is associated with increased polyamine synthesis, contributing to the deregulation of tumor metabolism and increased tumorigenesis [51,52]. When we investigated whether high-*MYC* tumors correlated with clinical parameters, we found patients with such tumors were in the overweight and obese categories (Table 2). Interestingly, the unfolded protein response (UPR), another hallmark of cancer that was enriched in our cohort's tumors (Supplementary Figure S4D,E), may be caused by obesity-induced endoplasmic reticulum (ER) stress [53,54]. MYC may regulate the UPR by increasing ER stress and directly regulating the transcription of UPR regulators [55]. In turn, the UPR may regulate MYC directly at the transcript level or through the promotion of MYC translation [55]. Therefore, obesity may be contributing to MYC activity and tumorigenesis in EOCRC in part through the UPR. In addition, obesity has been shown to increase intestinal *MYC* expression through the Wnt/β-catenin pathway [56]. The importance of identifying a molecular link between obesity and CRC is highlighted by a recent study that showed that obesity is strongly associated with an increased risk of EOCRC [57]. It is likely that a combination of dietary factors, increased gut inflammation, and the microbiome all contribute to EOCRC in all patients [58]; however, more work is needed to evaluate the relationship between obesity and genetics within EOCRC.

Copy number alterations (CNA) are often found at oncogenic loci throughout the genome [59]. Indeed, CNA at the 8q24.21 locus, which contains the *MYC* gene, is common in many cancers, including CRC [25]. Pan et al. reported *MYC* CNA and increased *MYC* gene expression in colorectal cancers from younger (<60 years old) patients [60]. However, Lee et al. reported that *MYC* amplification was not associated with the age of CRC patients [61]. Here, we found gains in *MYC* copy number (>2.5 copies) in 35% of tumors surveyed compared to controls. Importantly, we observed a correlation between *MYC* copy number and increased *MYC* expression. Previous estimates of low *MYC* copy

number amplification (>2 copies) range between 8–15% in CRC [61–63]. Therefore, in our EOCRC patient population, *MYC* CNA occurs more frequently than expected and is likely one mechanism that accounts for increased *MYC* expression within tumors. Due to our sample size, it is difficult to conclude whether *MYC* CNA is acting as a driver or passenger of carcinogenesis. No correlation was found between overall disease stage and *MYC* CNA, though we did find that patients with lymphovascular invasion were more likely to exhibit *MYC* CNA ($p = 0.028$, Supplementary Table S3). In addition, *MYC* CNA was found in older individuals ($p = 0.0007$), which may be supportive of amplification at the *MYC/PVT1* locus in tumor progression rather than initiation.

Prior work found that in 98% of cancers with MYC copy number increases, the amplified region includes the *PVT1* gene [25]. We found that all tumors with *MYC* copy number gains also displayed gains in *PVT1*, which supports the notion that the chromosomal 8q24 region harboring these genes is subjected to co-amplification in tumors. Other work has shown that *PVT1* cooperates with MYC to promote tumorigenesis [24–26]. In addition, a recent study has found that the *MYC/PVT1* locus is epigenetically regulated and may be subject to hypomethylation in CRC, which may also contribute to increased *MYC* expression [26]. In our study, we found that *PVT1* expression is increased in tumor samples and that it correlates with increased *MYC* expression. This finding suggests that while the *PVT1/MYC* region is not amplified in the majority of EOCRCs, *PVT1* may nonetheless serve as an important cooperating oncogene with *MYC* in tumors that develop in this patient population. The molecular mechanisms accounting for increased *MYC* and *PVT1* expression in most tumor samples in our study are currently unknown, and this line of work will be addressed in future studies.

In this study, we were able to detect a strong association between *MYC* levels and tumorigenesis despite the area of the colon from which the tumor was resected, arguing that *MYC* may be a critical factor in tumorigenesis. The association of high-*MYC* tumors and obesity may provide the beginnings of a molecular explanation for the rising incidence of EOCRC in developed countries, which also display increasing rates of obesity [64]. Though diagnostic testing can detect EOCRC early on, the American Cancer Society only recently began recommending that CRC screenings begin at age 45 instead of 50 [6]. Furthermore, there is a lack of compliance with CRC testing, especially in younger patients [65], highlighting a crucial need for EOCRC biomarkers to improve upon less invasive early detection options. A previous study suggested that MYC is among a group of genes whose epigenetic signature may be used as an early biomarker for detecting colorectal cancer [66]. Our work may indicate that *MYC* profiling is useful for the detection of EOCRC as well as LOCRC. Furthermore, we have shown that high and low-*MYC* tumors have distinct transcriptomes and thus, likely have distinct mechanisms of disease pathogenesis. Therefore, measuring MYC and identifying high, and low-MYC tumors could inform treatment options. More work is needed to define the molecular mechanisms in which MYC is deregulated in EOCRC and how MYC may contribute to EOCRC progression.

5. Limitations

As is the case with the majority of EOCRC transcriptome studies to date, ours also contains a relatively small sample size. Our samples were from different tissues throughout the colon, and we included patients with MSI, Lynch, and IBD-associated cancers, which are associated with distinct CRC molecular subtypes [67,68]. Full-thickness tissues were homogenized for sequencing, and thus, we cannot differentiate the transcriptomic signatures of different cell types. Furthermore, as there is no well-established model for EOCRC, we were unable to confirm molecular mechanisms and focus only on correlations within the data. Finally, as our patient population is composed of largely Caucasians, we are unable to address racial and ethnic disparities in our study.

References

1. Siegel, R.L.; Miller, K.D.; Fuchs, H.E.; Jemal, A. Cancer statistics, 2022. *CA Cancer J. Clin.* **2022**, *72*, 7–33. [CrossRef] [PubMed]
2. Hofseth, L.J.; Hebert, J.R.; Chanda, A.; Chen, H.; Love, B.L.; Pena, M.M.; Murphy, E.A.; Sajish, M.; Sheth, A.; Buckhaults, P.J.; et al. Early-onset colorectal cancer: Initial clues and current views. *Nat. Rev. Gastroenterol. Hepatol.* **2020**, *17*, 352–364. [CrossRef] [PubMed]
3. Patel, S.G.; Ahnen, D.J. Colorectal Cancer in the Young. *Curr. Gastroenterol. Rep.* **2018**, *20*, 15. [CrossRef] [PubMed]
4. Cercek, A.; Chatila, W.K.; Yaeger, R.; Walch, H.; Fernandes, G.D.S.; Krishnan, A.; Palmaira, L.; Maio, A.; Kemel, Y.; Srinivasan, P.; et al. A Comprehensive Comparison of Early-Onset and Average-Onset Colorectal Cancers. *JNCI J. Natl. Cancer Inst.* **2021**, *113*, 1683–1692. [CrossRef]
5. Cavestro, G.M.; Mannucci, A.; Zuppardo, R.A.; Di Leo, M.; Stoffel, E.; Tonon, G. Early onset sporadic colorectal cancer: Worrisome trends and oncogenic features. *Dig. Liver Dis.* **2018**, *50*, 521–532. [CrossRef]
6. Wender, R.C. Colorectal cancer screening should begin at 45. *J. Gastroenterol. Hepatol.* **2020**, *35*, 1461–1463. [CrossRef]
7. Mauri, G.; Sartore-Bianchi, A.; Russo, A.G.; Marsoni, S.; Bardelli, A.; Siena, S. Early-onset colorectal cancer in young individuals. *Mol. Oncol.* **2019**, *13*, 109–131. [CrossRef]
8. Keum, N.; Giovannucci, E. Global burden of colorectal cancer: Emerging trends, risk factors and prevention strategies. *Nat. Rev. Gastroenterol. Hepatol.* **2019**, *16*, 713–732. [CrossRef]
9. Strum, W.B.; Boland, C.R. Clinical and Genetic Characteristics of Colorectal Cancer in Persons under 50 Years of Age: A Review. *Dig. Dis. Sci.* **2019**, *64*, 3059–3065. [CrossRef]
10. Álvaro, E.; Cano, J.M.; García, J.L.; Brandáriz, L.; Olmedillas-López, S.; Arriba, M.; Rueda, D.; Rodríguez, Y.; Cañete, Á.; Arribas, J.; et al. Clinical and Molecular Comparative Study of Colorectal Cancer Based on Age-of-onset and Tumor Location: Two Main Criteria for Subclassifying Colorectal Cancer. *Int. J. Mol. Sci.* **2019**, *20*, 968. [CrossRef]
11. Perea, J.; Rueda, D.; Canal, A.; Rodríguez, Y.; Álvaro, E.; Osorio, I.; Alegre, C.; Rivera, B.; Martínez, J.; Benítez, J.; et al. Age at onset should be a major criterion for subclassification of colorectal cancer. *J. Mol. Diagn.* **2014**, *16*, 116–126. [CrossRef] [PubMed]
12. Lee, W.; Wang, Z.; Saffern, M.; Jun, T.; Huang, K.L. Genomic and molecular features distinguish young adult cancer from later-onset cancer. *Cell Rep.* **2021**, *37*, 110005. [CrossRef] [PubMed]
13. Zaborowski, A.M.; Abdile, A.; Adamina, M.; Aigner, F.; D'Allens, L.; Allmer, C.; Álvarez, A.; Anula, R.; Andric, M.; Atallah, S.; et al. Characteristics of Early-Onset vs Late-Onset Colorectal Cancer: A Review. *JAMA Surg.* **2021**, *156*, 865–874. [CrossRef] [PubMed]
14. Chen, H.; Liu, H.; Qing, G. Targeting oncogenic Myc as a strategy for cancer treatment. *Signal Transduct. Target. Ther.* **2018**, *3*, 5. [CrossRef]

15. Dang, C.V.; O'Donnell, K.A.; Zeller, K.I.; Nguyen, T.; Osthus, R.C.; Li, F. The c-Myc target gene network. *Semin. Cancer Biol.* **2006**, *16*, 253–264. [CrossRef] [PubMed]

16. Adhikary, S.; Eilers, M. Transcriptional regulation and transformation by Myc proteins. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 635–645. [CrossRef] [PubMed]

17. Pelengaris, S.; Khan, M.; Evan, G. c-MYC: More than just a matter of life and death. *Nat. Rev. Cancer* **2002**, *2*, 764–776. [CrossRef] [PubMed]

18. Carroll, P.A.; Freie, B.W.; Mathsyaraja, H.; Eisenman, R.N. The MYC transcription factor network: Balancing metabolism, proliferation and oncogenesis. *Front. Med.* **2018**, *12*, 412–425. [CrossRef]

19. Gurel, B.; Iwata, T.; Koh, C.M.; Jenkins, R.B.; Lan, F.; Van Dang, C.; Hicks, J.L.; Morgan, J.; Cornish, T.C.; Sutcliffe, S.; et al. Nuclear MYC protein overexpression is an early alteration in human prostate carcinogenesis. *Mod. Pathol.* **2008**, *21*, 1156–1167. [CrossRef]

20. Choi, W.; Kim, J.; Park, J.; Lee, D.-H.; Hwang, D.; Kim, J.-H.; Ashktorab, H.; Smoot, D.; Kim, S.-Y.; Choi, C.; et al. YAP/TAZ Initiates Gastric Tumorigenesis via Upregulation of MYC. *Cancer Res.* **2018**, *78*, 3306–3320. [CrossRef]

21. Sollazzo, M.; Genchi, C.; Paglia, S.; Di Giacomo, S.; Pession, A.; de Biase, D.; Grifoni, D. High MYC Levels Favour Multifocal Carcinogenesis. *Front. Genet.* **2018**, *9*, 612. [CrossRef] [PubMed]

22. Meyer, N.; Penn, L.Z. Reflecting on 25 years with MYC. *Nat. Rev. Cancer* **2008**, *8*, 976–990. [CrossRef]

23. Kalkat, M.; De Melo, J.; Hickman, K.A.; Lourenco, C.; Redel, C.; Resetca, D.; Tamachi, A.; Tu, W.B.; Penn, L.Z. MYC Deregulation in Primary Human Cancers. *Genes* **2017**, *8*, 151. [CrossRef] [PubMed]

24. Boloix, A.; Masanas, M.; Jiménez, C.; Antonelli, R.; Soriano, A.; Roma, J.; De Toledo, J.S.; Gallego, S.; Segura, M.F. Long Non-coding RNA PVT1 as a Prognostic and Therapeutic Target in Pediatric Cancer. *Front. Oncol.* **2019**, *9*, 1173. [CrossRef] [PubMed]

25. Tseng, Y.-Y.; Moriarity, B.S.; Gong, W.; Akiyama, R.; Tiwari, A.; Kawakami, H.; Ronning, P.; Reuland, B.; Guenther, K.; Beadnell, T.C.; et al. PVT1 dependence in cancer with MYC copy-number increase. *Nature* **2014**, *512*, 82–86. [CrossRef] [PubMed]

26. Shigeyasu, K.; Toden, S.; Ozawa, T.; Matsuyama, T.; Nagasaka, T.; Ishikawa, T.; Sahoo, D.; Ghosh, P.; Uetake, H.; Fujiwara, T.; et al. The PVT1 lncRNA is a novel epigenetic enhancer of MYC, and a promising risk-stratification biomarker in colorectal cancer. *Mol. Cancer* **2020**, *19*, 155. [CrossRef]

27. Rennoll, S.A.; Eshelman, M.A.; Raup-Konsavage, W.M.; Kawasawa, Y.I.; Yochum, G.S. The MYC 3' Wnt-Responsive Element Drives Oncogenic MYC Expression in Human Colorectal Cancer Cells. *Cancers* **2016**, *8*, 52. [CrossRef]

28. Zeng, W.Z.D.; Glicksberg, B.S.; Li, Y.; Chen, B. Selecting precise reference normal tissue samples for cancer research using a deep learning approach. *BMC Med. Genom.* **2019**, *12* (Suppl. 1), 21. [CrossRef]

29. Gross, A.M.; Kreisberg, J.F.; Ideker, T. Analysis of Matched Tumor and Normal Profiles Reveals Common Transcriptional and Epigenetic Signals Shared across Cancer Types. *PLoS ONE* **2015**, *10*, e0142618. [CrossRef]

30. Schieffer, K.M.; Choi, C.S.; Emrich, S.; Harris, L.; Deiling, S.; Karamchandani, D.M.; Salzberg, A.; Kawasawa, Y.I.; Yochum, G.S.; Koltun, W.A. RNA-seq implicates deregulation of the immune system in the pathogenesis of diverticulitis. *Am. J. Physiol.-Gastrointest. Liver Physiol.* **2017**, *313*, G277–G284. [CrossRef]

31. Eshelman, M.A.; Jeganathan, N.A.; Schieffer, K.M.; Kline, B.P.; Mendenhall, M.; Deiling, S.; Harris, L.; Koltun, W.A.; Yochum, G.S. Elevated Colonic Mucin Expression Correlates with Extended Time to Surgery for Ulcerative Colitis Patients. *J. Gastrointest. Liver Dis.* **2019**, *28*, 405–413. [CrossRef] [PubMed]

32. Schieffer, K.M.; Kline, B.P.; Harris, L.R.; Deiling, S.; Koltun, W.A.; Yochum, G.S. A differential host response to viral infection defines a subset of earlier-onset diverticulitis patients. *J. Gastrointest. Liver Dis.* **2018**, *27*, 249–255. [CrossRef] [PubMed]

33. Dobin, A.; Davis, C.A.; Schlesinger, F.; Drenkow, J.; Zaleski, C.; Jha, S.; Batut, P.; Chaisson, M.; Gingeras, T.R. STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **2013**, *29*, 15–21. [CrossRef] [PubMed]

34. Anders, S.; Pyl, P.T.; Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **2015**, *31*, 166–169. [CrossRef] [PubMed]

35. Love, M.I.; Huber, W.; Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **2014**, *15*, 550. [CrossRef]

36. Subramanian, A.; Tamayo, P.; Mootha, V.K.; Mukherjee, S.; Ebert, B.L.; Gillette, M.A.; Paulovich, A.; Pomeroy, S.L.; Golub, T.R.; Lander, E.S.; et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 15545–15550. [CrossRef]

37. Wu, T.; Hu, E.; Xu, S.; Chen, M.; Guo, P.; Dai, Z.; Feng, T.; Zhou, L.; Tang, W.; Zhan, L.; et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation* **2021**, *2*, 100141. [CrossRef]

38. Yu, G.; Wang, L.G.; Han, Y.; He, Q.Y. clusterProfiler: An R package for comparing biological themes among gene clusters. *Omics* **2012**, *16*, 284–287. [CrossRef]

39. Aran, D.; Hu, Z.; Butte, A.J. xCell: Digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol.* **2017**, *18*, 220. [CrossRef]

40. Villanueva, R.A.M.; Chen, Z.J. ggplot2: Elegant Graphics for Data Analysis (2nd ed.). *Meas. Interdiscip. Res. Perspect.* **2019**, *17*, 160–167. [CrossRef]

41. Colaprico, A.; Silva, T.C.; Olsen, C.; Garofano, L.; Cava, C.; Garolini, D.; Sabedot, T.S.; Malta, T.M.; Pagnotta, S.M.; Castiglioni, I.; et al. TCGAbiolinks: An R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res.* **2016**, *44*, e71. [CrossRef] [PubMed]

42. Liberzon, A.; Birger, C.; Thorvaldsdottir, H.; Ghandi, M.; Mesirov, J.P.; Tamayo, P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* **2015**, *1*, 417–425. [CrossRef] [PubMed]

43. Yan, H.H.N.; Siu, H.C.; Ho, S.L.; Yue, S.S.K.; Gao, Y.; Tsui, W.Y.; Chan, D.; Chan, A.S.; Wong, J.W.H.; Man, A.H.Y.; et al. Organoid cultures of early-onset colorectal cancers reveal distinct and rare genetic profiles. *Gut* **2020**, *69*, 2165–2179. [CrossRef] [PubMed]

44. Clevers, H. Wnt/beta-catenin signaling in development and disease. *Cell* **2006**, *127*, 469–480. [CrossRef]

45. Vita, M.; Henriksson, M. The Myc oncoprotein as a therapeutic target for human cancer. *Semin. Cancer Biol.* **2006**, *16*, 318–330. [CrossRef]

46. Hong, Y.; Ho, K.S.; Eu, K.W.; Cheah, P.Y. A susceptibility gene set for early onset colorectal cancer that integrates diverse signaling pathways: Implication for tumorigenesis. *Clin. Cancer Res.* **2007**, *13*, 1107–1114. [CrossRef]

47. Fearon, E.R. Molecular genetics of colorectal cancer. *Annu. Rev. Pathol. Mech. Dis.* **2011**, *6*, 479–507. [CrossRef]

48. Xu, T.; Zhang, Y.; Zhang, J.; Qi, C.; Liu, D.; Wang, Z.; Li, Y.; Ji, C.; Li, J.; Lin, X.; et al. Germline Profiling and Molecular Characterization of Early Onset Metastatic Colorectal Cancer. *Front. Oncol.* **2020**, *10*, 568911. [CrossRef]

49. Lieu, C.H.; Golemis, E.A.; Serebriiskii, I.G.; Newberg, J.; Hemmerich, A.; Connelly, C.; Messersmith, W.A.; Eng, C.; Eckhardt, S.G.; Frampton, G.; et al. Comprehensive Genomic Landscapes in Early and Later Onset Colorectal Cancer. *Clin. Cancer Res.* **2019**, *25*, 5852–5858. [CrossRef]

50. Bachmann, A.S.; Geerts, D. Polyamine synthesis as a target of MYC oncogenes. *J. Biol. Chem.* **2018**, *293*, 18757–18769. [CrossRef]

51. Snezhkina, A.V.; Krasnov, G.S.; Lipatova, A.V.; Sadritdinova, A.F.; Kardymon, O.L.; Fedorova, M.S.; Melnikova, N.V.; Stepanov, O.A.; Zaretsky, A.R.; Kaprin, A.D.; et al. The Dysregulation of Polyamine Metabolism in Colorectal Cancer Is Associated with Overexpression of c-Myc and C/EBPβ rather than Enterotoxigenic Bacteroides fragilis Infection. *Oxidative Med. Cell. Longev.* **2016**, *2016*, 2353560. [CrossRef] [PubMed]

52. Guo, Y.; Ye, Q.; Deng, P.; Cao, Y.; He, D.; Zhou, Z.; Wang, C.; Zaytseva, Y.Y.; Schwartz, C.E.; Lee, E.Y.; et al. Spermine synthase and MYC cooperate to maintain colorectal cancer cell survival by repressing Bim expression. *Nat. Commun.* **2020**, *11*, 3243. [CrossRef] [PubMed]

53. Fu, S.; Yang, L.; Li, P.; Hofmann, O.; Dicker, L.; Hide, W.; Lin, X.; Watkins, S.M.; Ivanov, A.R.; Hotamisligil, G.S. Aberrant lipid metabolism disrupts calcium homeostasis causing liver endoplasmic reticulum stress in obesity. *Nature* **2011**, *473*, 528–531. [CrossRef]

54. Ron, D.; Walter, P. Signal integration in the endoplasmic reticulum unfolded protein response. *Nat. Rev. Mol. Cell Biol.* **2007**, *8*, 519–529. [CrossRef]

55. Zhang, T.; Li, N.; Sun, C.; Jin, Y.; Sheng, X. MYC and the unfolded protein response in cancer: Synthetic lethal partners in crime? *EMBO Mol. Med.* **2020**, *12*, e11845. [CrossRef] [PubMed]

56. Liu, Z.; Brooks, R.S.; Ciappio, E.D.; Kim, S.J.; Crott, J.W.; Bennett, G.; Greenberg, A.S.; Mason, J.B. Diet-induced obesity elevates colonic TNF-α in mice and is accompanied by an activation of Wnt signaling: A mechanism for obesity-associated colorectal cancer. *J. Nutr. Biochem.* **2012**, *23*, 1207–1213. [CrossRef]

57. Li, H.; Boakye, D.; Chen, X.; Hoffmeister, M.; Brenner, H. Association of Body Mass Index With Risk of Early-Onset Colorectal Cancer: Systematic Review and Meta-Analysis. *Am. J. Gastroenterol.* **2021**, *116*, 2173–2183. [CrossRef]

58. Akimoto, N.; Ugai, T.; Zhong, R.; Hamada, T.; Fujiyoshi, K.; Giannakis, M.; Wu, K.; Cao, Y.; Ng, K.; Ogino, S. Rising incidence of early-onset colorectal cancer—A call to action. *Nat. Rev. Clin. Oncol.* **2021**, *18*, 230–243. [CrossRef]

59. Zack, T.I.; Schumacher, S.E.; Carter, S.L.; Cherniack, A.D.; Saksena, G.; Tabak, B.; Lawrence, M.S.; Zhang, C.Z.; Wala, J.; Mermel, C.H.; et al. Pan-cancer patterns of somatic copy number alteration. *Nat. Genet.* **2013**, *45*, 1134–1140. [CrossRef]

60. Pan, W.; Wang, W.; Huang, J.; Lu, K.; Huang, S.; Jiang, D.; Bu, D.; Liu, J.; Jing, H.; Yao, J.; et al. The prognostic role of c-MYC amplification in schistosomiasis-associated colorectal cancer. *Jpn. J. Clin. Oncol.* **2020**, *50*, 446–455. [CrossRef]

61. Lee, K.S.; Kwak, Y.; Nam, K.H.; Kim, D.-W.; Kang, S.-B.; Choe, G.; Kim, W.H.; Lee, H.S. c-MYC Copy-Number Gain Is an Independent Prognostic Factor in Patients with Colorectal Cancer. *PLoS ONE* **2015**, *10*, e0139727. [CrossRef] [PubMed]

62. Al-Kuraya, K.; Novotny, H.; Bavi, P.; Siraj, A.K.; Uddin, S.; Ezzat, A.; Sanea, N.A.; Al-Dayel, F.; Al-Mana, H.; Sheikh, S.S.; et al. HER2, TOP2A, CCND1, EGFR and C-MYC oncogene amplification in colorectal cancer. *J. Clin. Pathol.* **2007**, *60*, 768–772. [CrossRef] [PubMed]

63. Kwak, Y.; Yun, S.; Nam, S.K.; Seo, A.N.; Lee, K.S.; Shin, E.; Oh, H.-K.; Kim, D.W.; Kang, S.B.; Kim, W.H.; et al. Comparative analysis of the EGFR, HER2, c-MYC, and MET variations in colorectal cancer determined by three different measures: Gene copy number gain, amplification status and the 2013 ASCO/CAP guideline criterion for HER2 testing of breast cancer. *J. Transl. Med.* **2017**, *15*, 167. [CrossRef] [PubMed]

64. Siegel, R.L.; Torre, L.A.; Soerjomataram, I.; Hayes, R.B.; Bray, F.; Weber, T.K.; Jemal, A. Global patterns and trends in colorectal cancer incidence in young adults. *Gut* **2019**, *68*, 2179–2185. [CrossRef]

65. Stanich, P.P.; Pelstring, K.R.; Hampel, H.; Pearlman, R. A High Percentage of Early-age Onset Colorectal Cancer Is Potentially Preventable. *Gastroenterology* **2021**, *160*, 1850–1852. [CrossRef]

66. Kel, A.; Boyarskikh, U.; Stegmaier, P.; Leskov, L.S.; Sokolov, A.V.; Yevshin, I.; Mandrik, N.; Stelmashenko, D.; Koschmann, J.; Kel-Margoulis, O.; et al. Walking pathways with positive feedback loops reveal DNA methylation biomarkers of colorectal cancer. *BMC Bioinform.* **2019**, *20* (Suppl. 4), 119. [CrossRef]

67. Müller, M.F.; Ibrahim, A.E.K.; Arends, M.J. Molecular pathological classification of colorectal cancer. *Virchows Arch.* **2016**, *469*, 125–134. [CrossRef]

68. Rajamäki, K.; Taira, A.; Katainen, R.; Välimäki, N.; Kuosmanen, A.; Plaketti, R.-M.; Seppälä, T.T.; Ahtiainen, M.; Wirta, E.-V.; Vartiainen, E.; et al. Genetic and Epigenetic Characteristics of Inflammatory Bowel Disease-Associated Colorectal Cancer. *Gastroenterology* **2021**, *161*, 592–607. [CrossRef]