Article

Insights into Early Onset Colorectal Cancer through Analysis of Normal Colon Organoids of Familial Adenomatous Polyposis Patients

Matthew A. Devall [1,2] **, Stephen Eaton** [1,2] **, Mourad W. Ali** [1] **, Steven M. Powell** [3] **, Li Li** [2,4,]*
and Graham Casey [1,4,5,]*

1   Center for Public Health Genomics, University of Virginia, Charlottesville, VA 22908, USA
2   Department of Family Medicine, University of Virginia, Charlottesville, VA 22903, USA
3   Digestive Health Center, University of Virginia, Charlottesville, VA 22903, USA
4   Comprehensive Cancer Center, University of Virginia, Charlottesville, VA 22911, USA
5   Department of Public Health Sciences, University of Virginia, Charlottesville, VA 22908, USA
*   Correspondence: ll8nv@virginia.edu (L.L.); gc8r@virginia.edu (G.C.)

Simple Summary: The increase in early onset colorectal cancer (EOCRC) rates places a new, heavy burden on a younger population. Current studies in EOCRC are often affected by factors such as increased cellular heterogeneity, which mask biological signal. We overcome this using the colon organoid model, which is composed of cells from which CRC is believed to originate. We perform gene expression analysis on colon organoids derived from healthy and familial adenomatous polyposis patients, who are genetically predisposed to develop CRC at a young age, to identify differences that may be used for early monitoring of CRC development. We contextualize our findings in the framework of existing CRC and EOCRC data. We also show that ethanol, a CRC risk factor, may play an important role in further driving aberrant gene expression at EOCRC relevant genes. Functional analysis of these genes may shed new insight into EOCRC.

Abstract: Early onset colorectal cancer (EOCRC) rates have increased in recent decades. While lowering the recommended age for routine colonoscopies to 45 may reduce this burden, such measures do not address those who develop CRC before that age. Additional measures are needed to identify individuals at-risk for CRC. To better define transcriptomic events that precede the development of CRC, we performed RNA-sequencing analysis in colon organoids derived from seven healthy and six familial adenomatous polyposis (FAP) patients. This led to the identification of 2635 significant differentially expressed genes (FDR < 0.05). Through secondary analysis of publicly available datasets, we found that these genes were enriched for significant genes also present in FAP CRC and non-hereditary CRC datasets, including a subset that were unique to EOCRC. By exposing FAP colon organoids to a three-day ethanol treatment, we found that two EOCRC-relevant genes were also targets of CRC related lifestyle factors. Our data provides unique insight into the potential, early mechanisms of CRC development in colon epithelial cells, which may provide biomarkers for patient monitoring. We also show how modifiable lifestyle factors may further alter genes relevant to EOCRC, adding weight to the hypothesis that such factors represent an important contributor to increased EOCRC incidence.

Keywords: familial adenomatous polyposis; early onset colorectal cancer; colon organoid; RNA-seq

1. Introduction

Colorectal cancer (CRC) is a complex, heterogenous disease with a strong genetic component [1–3]. Overall, the incidence of CRC has decreased steadily since the mid 2000's. A decline that has largely been attributed to improved adherence to CRC screening programs [4]. In contrast, there has been an alarming increase in the incidence of early onset

CRC (EOCRC) (defined as <50 years old) [4,5]. The initial increase in EOCRC in the US was primarily driven by rectal cancer, but since 2012, incidence of cancers of both colon and rectum have been increasing at similar rates [6]. This increase in EOCRC places a significant burden on a younger population, where screening has not been routinely recommended until recently [7]. While it is now recommended that individuals begin by the age of 45, this does little to aid those who develop CRC at a younger age, where increased incidence has also been reported [8].

The vast majority of EOCRCs are believed to be driven by non-hereditary forms of CRC [9]. While the causes of these 'sporadic' EOCRC remain poorly understood, generational shifts in environment, lifestyle, and dietary habits have been strongly implicated [10]. For example, a recent study has provided evidence that moderate to high consumption of alcohol is associated with a greater risk of EOCRC than late onset CRC [11]. However, approximately one fifth of patients with EOCRC are affected by a hereditary cancer syndrome [4], which is about four times greater than the rates of these syndromes in CRC across all ages. Familial adenomatous polyposis (FAP) is a hereditary syndrome that leads to colon cancer. It is driven by inherited or de novo mutations in the tumor suppressor gene APC regulator of WNT signaling pathway (*APC*) [12]. Patients with FAP typically develop hundreds of adenomatous polyps in their late teens or early twenties [13]. If left untreated, the development of CRC is almost inevitable and develops at an average age of 39 [12,14]. While somewhat rare in the general population (1/7500–1/10,000) [13], FAP is the second most common, hereditary CRC syndrome [15], accounting for approximately one percent of all CRC cases [12]. Discovery of early aberrant gene expression patterns in uninvolved colons of FAP patients has the potential to provide critical insight to improve our understanding of early events of EOCRC oncogenesis.

A common limitation in EOCRC research is that the majority of studies have been performed in tumors; for example, within the context of identifying gene expression differences between tumor versus normal-adjacent tissue (NAT) [16]. While this has improved our overall understanding of molecular changes in EOCRC and CRC tumors as a whole, tumor and cellular heterogeneity, increased somatic mutation burden and gross alterations in copy number landscapes in CRC tumors present obstacles for the interpretation of gene expression differences that drive EOCRC. While the incident of EOCRC is rising, EOCRC remains a relatively rare disease and data from few studies on EOCRC tumors have made publicly available, limiting further advancement in the field. Importantly, there is also a lack of well-defined models for the study of EOCRC. Given these limitations, we sought to make use of the colon organoid system (which present with reduced cellular and tumoral heterogeneity) derived from "normal appearing" colon biopsies of FAP patients (i.e., biopsies taken from sites adjacent to polyps during surveillance colonoscopies from patients without known, existing CRC but with extensive polyposis) in an attempt to improve understanding of early events of EOCRC and CRC as a whole.

Colon organoids represent a model of the colon stem cell niche, where CRC is believed to originate [17]. Previously, our group has used this model to investigate how differences in CRC environmental risk factors may alter gene expression of genes implicated in CRC [18–22]. In this study, we have applied the colon organoid model to identify gene expression differences between FAP subjects and healthy controls. Given the limited available models for EOCRC, we hypothesize that the FAP colon organoid model may be used to gain unique insight into EOCRC and CRC development. Specifically, we aimed to identify CRC-related events in at-risk subjects without existing CRC. Identification of these early biological markers could lead to improved early detection of CRC, including in younger population.

2. Materials and Methods

2.1. Patient Selection and Biopsy Collection

Healthy subjects (*n* = 7) undergoing standard of care colonoscopy were recruited at the University of Virginia alongside FAP (*n* = 6) who were undergoing surveillance colonoscopies. All healthy subjects included in the study presented with 3 or fewer polyps and no personal or immediate family history of CRC. FAP patients were defined by clinical presentation and/or genetic mutation. FAP and healthy colon organoids displayed no statistically significant differences for age, biological sex or smoking status, though two of the seven healthy subjects were current or previous smokers. All subjects were self-reported to be non-Hispanic White.

2.2. Organoid Establishment and Passaging

This study included organoids derived from the left colon for both healthy controls and FAP patients. Biopsy samples were first washed 3 times with DPBS at room temperature. Following this, 10 mL of 9 mM EDTA in room temperature DPBS was added to each biopsy. Biopsies were incubated in DPBS/EDTA for 20 min at room temperature. During this period, samples were manually inverted every 2–3 min. To avoid crypt adhesion, all tubes, tips and pipets were conditioned with wash media (DMEM/F-12, 10% FBS, 10 mM HEPES, 2 mM L-glutamine, 1X Pen-Strep (100 U/mL penicillin, 100 µg/mL streptomycin), 1X Glutamax) during the incubation period. Following incubation. DPBS/EDTA was removed. 10 mL of DPBS without EDTA was added. This volume was then mixed 8–10 times by manual pipetting. After the tissue settled, any supernatant was collected in a 15 mL tube. This process was repeated a total of 3 times. The supernatant was spun at 1200 rpm at 4 °C for 5 min. All but 2 mL of supernatant from each 15 mL tube was then removed and combined in a clean 50 mL tube. Care was taken not to disturb the pellet. Using a conditioned 5 mL pipet, the remaining 2 mL from each tube was then combined into one 15 mL tube, which was then spun at 1200 rpm at 4 °C for 5 min. Following this, the supernatant was removed. 10 µL of complex media (45% Wash Media, 50% L-WRN conditioned media, 10 nM Gastrin, 10 µM Y27632, 1X B27 Supplement, 1X N2 Supplement, 1 mM n-Acetylcysteine, 50 ng/µL EGF, 10 mM Nicotinamide, 500 nM A83, 10 µM SB202190) was added. To reduce the clumping of crypts, the cell solution was mixed through manual pipetting. 200 µL of Matrigel was then added and mixed and 35–50 µL of each sample was plated at the center of wells of a 48-well plate. Matrigel containing cells was then incubated at 37 °C for 15 min. Following this, 300–500 µL of complex media was added and samples were returned to the incubator. Organoids were checked and fed after 24 h. Upon establishment, organoids were subsequently passaged as needed (every 3–5 days) and then frozen. A subset of FAP and healthy colon organoids were imaged during routine passaging using Lumenera Infinity2-2C 2.0 Megapixel CCD Color Camera (Lumenera, Ottawa, Canada, cat. #95107) and Infinity Analyze software version 6.5.5 (Lumenera) at 100× magnification.

2.3. RNA Harvesting, Library Preparation, Sequencing and Pre-Processing of Colon Organoid RNA-Seq Data

The NucleoSpin RNA Mini kit (Macherey-Nagel, Düren, Germany) was used to extract total RNA and RNA integrity numbers (RIN) were measured using the Agilent 4200 Tapestation (Santa Clara, United States, cat. G2991BA). RIN values of samples considered within this study ranged from 9.3–10. No significant differences in RIN values were observed between the two phenotypes. Libraries preparation was carried out using the Illumina Stranded mRNA Prep kit. Sequencing of RNA libraries was performed in accordance with standard Illumina protocols within the Genomics Core Facility of the Center for Advanced Genomics Technology (Icahn School of Medicine at Mount Sinai, New York). Libraries were pooled and 100 bp, paired-end reads were sequenced on a NovaSeq 6000. FASTQ files were aligned to GENCODE v29 [23] using STAR (v2.6.1d) [24] and RSEM (v1.3.1) [25].

2.4. Collection of Publicly Available Data

Publicly available data in the form of raw count matrices were downloaded for our secondary analysis. Matched tumor versus NAT data of human FAP patients was downloaded from Gene Expression Omnibus [26], accession number, GSE153385 [27]. Similarly, raw count files from The Cancer Genome Atlas Colon Adenocarcinoma cohort (TCGA-COAD) [16] were downloaded using the R package, TCGAbiolinks [28]. For analysis of TCGA RNA-seq data, subjects were split into two categories: those who developed CRC before the age of 50 (*n* = 5 pairs; 40–48 y/o; herein TCGA-EOCRC) and those who developed CRC at or after the age of 60 (*n* = 34 pairs; 60–90 y/o; herein TCGA-60+). Data on subjects diagnosed between the ages of 50–59 were excluded from our study.

2.5. Statistical Analysis of RNA-Seq

Gene expression analysis was carried out in R (version 4.1.1) [29]. For novel, FAP versus healthy colon organoid data, gene abundance estimates from RSEM [25] were converted to gene counts using tximport [30] and differential expression between groups was tested by fitting negative binomial generalized linear regression models for each gene in DESeq2 [31]. Genes were considered significant if they survived a 5% adjusted false discovery rate (FDR) in DESeq2 [31]. Regression models on phenotype were corrected for biological sex and age. Pathway analysis was performed by uploading significant (FDR < 0.05) DEGs to STRING under default settings [32]. Pathway reporting was limited to Gene Ontology (GO) [33] and Kyoto Encyclopedia of Genes and Genomes (KEGG) [34] enrichment terms. For data visualization, the lfcShrink() function of DESeq2 was performed while specifying apeglm [35]. A volcano plot was then generated using the EnhancedVolcano package on these adjusted results [36]. For matched, tumor versus NAT data, a paired regression model was considered in DESeq2 [31] while using sample pairing as a blocking factor to adjust for inter-individual differences.

2.6. Ethanol Treatment of FAP Colon Organoids

Ethanol exposures of FAP colon organoids (*n* = 6) were conducted as previously described [19,20]. FAP colon organoid lines were grown in 48 well plates at a density of approximately 10^5 cells per well. Organoids were treated with fresh growth media plus ethyl alcohol (Sigma Aldrich, St. Louis, United States, cat. E7023) (2 µL per 1 mL of growth media). Organoids within the control group were instead treated with fresh media plus 2 µL cell culture grade distilled water as a control. To replace evaporating ethanol, media was replaced every 24 h for three days with fresh ethanol or water control. For consistency, this dose and timeframe were chosen specifically to determine whether the effects of ethanol previously described in normal colon organoids [19,20], could also be identified in colon organoids derived from FAP patients. In previous studies, this dose was chosen to reflect the circulating alcohol levels in blood of alcoholics and was similar to the dose of blood alcohol levels determined through in vivo studies of rats consuming between 2–4 drinks [37,38].

2.7. Quantitative PCR

For technical validation of selected significant findings in colon organoids, existing RNA was considered. For replication of findings in matched left colon biopsies (*n* = 4), RNA was extracted using an AllPrep RNA kit (Qiagen, Hilden, Germany, cat. 80284), followed by DNAse treatment using an RNA Clean and Concentrator kit (Zymo, Irivine, United States, cat. R1080). The following gene assays were considered for technical validation and replication of findings in colon organoids and biopsies: NKD inhibitor of WNT signaling pathway 1 (*NKD1*; Hs01548773_m1), EYA transcriptional coactivator and phosphatase 2 (*EYA2*; Hs00193347_m1), early growth response 2 (*EGR2*; Hs00166165_m1), *Ependymin related 1 (EPDR1*; Hs01556067_m1) and IGF like family member 1 (*IGFL1*; Hs01651089_g1). A subset of genes were also analyzed in six FAP colon organoids exposed to ethanol (DNA damage regulated autophagy modulator 1 (*DRAM1*; Hs01022842_m1), growth differenti-

ation factor 11 (*GDF11*; Hs00195156_m1), Notum, palmitoleoyl-protein carboxylesterase (*NOTUM*; Hs00394510_m1) and pleckstrin homology domain containing B1 (*PLEKHB1*; Hs00943921_m1)) for a period of three days. For each analysis, cDNA was synthesized from 2 μg of total RNA using the High-Capacity Reverse Transcriptase cDNA kit (Thermo Fisher) and quantitative real time polymerase chain reaction (qPCR) was performed using the TaqMan Gene Expression Master Mix (Thermo Fisher, Waltham, MA, USA, cat. 4444557) on selected TaqMan assays. Hypoxanthine phosphoribosyltransferase 1 (*HPRT1*; Hs02800695_m1) was used as an internal control gene. The PCR reactions were all ran in duplicate, and were subsequently analyzed using QuantStudio 5 (Thermo Fisher). Reactions were normalized using the control gene *HPRT1*, and calculations were performed according to the 2^{-ddCT} method. For the analysis of expression of FAP biopsy samples (*n* = 2), log fold differences were estimated for each gene in each FAP sample, by comparing the resultant values of that gene to that of the average of the resultant from colon biopsies of two healthy controls. For technical replication of findings, data were analyzed for statistical differences using a linear regression model on log normalized values while adjusting for age and gender.

3. Results

### 3.1. Defining the Transcriptomic Landscape of Colon Organoids of FAP Patients versus Those of Healthy Subjects

We compared gene expression of colon organoids generated from FAP patients (*n* = 6) and healthy subjects (*n* = 7). No significant differences for age, sex or smoking status were identified between the two phenotypes (Supplementary Table S1). We generated RNA-seq data on colon organoids from all subjects. An average mapping efficiency of 80.63% was observed, which led to an average of 33.26 million uniquely mapped reads per sample (Supplementary Table S2). Regression analysis was carried out in DESeq2 [31], while accounting for age at the time of colonoscopy and biological sex. This led to the identification of 2635 significant DEGs (FDR < 0.05), 59.66% of which displayed increased expression in colon organoids of FAP versus healthy subjects (Figure 1). Differentially methylated regions previously shown to be unique to EOCRC tumors corresponded to a subset of 23 of these genes (Figure 2) [39].

Pathway enrichment analysis was performed on all DEGs identified in our analysis using the STRING database [32]. Here, we noted that our DEG list was enriched for protein–protein interactions ($p < 1.0 \times 10^{-16}$), adding weight to their biological relatedness. DEGs were enriched for 104 gene ontology terms [33,34], as well as three KEGG terms: "Wnt signaling pathway" (FDR = 5.30×10^{-5}), "Hippo signaling pathway (FDR = 3.80×10^{-3}) and "mitophagy—animal" (FDR = 0.016) (Supplementary Table S3). Extracellular Wnt sensitivity and/or Wnt signaling has previously been suggested to be important for earlier onset of CRC development in other studies [40,41].

3.2. Relationship of DNA Methylation and Differential Expression in FAP Colon Organoids

Previously, we performed DNA methylation analysis (Illumina Infinium MethylationEPIC Kit) on a small (*n* = 23), largely overlapping (84.62%) cohort of FAP and healthy colon organoids [42]. We identified 358 FDR corrected, differentially methylated regions between FAP and healthy colon organoids, corresponding to 439 unique genes. Of these, 36 genes were found to also be significantly different in our current RNA-seq analysis (Supplementary Table S4), providing further evidence for an important role of these genes in FAP epithelial cells.
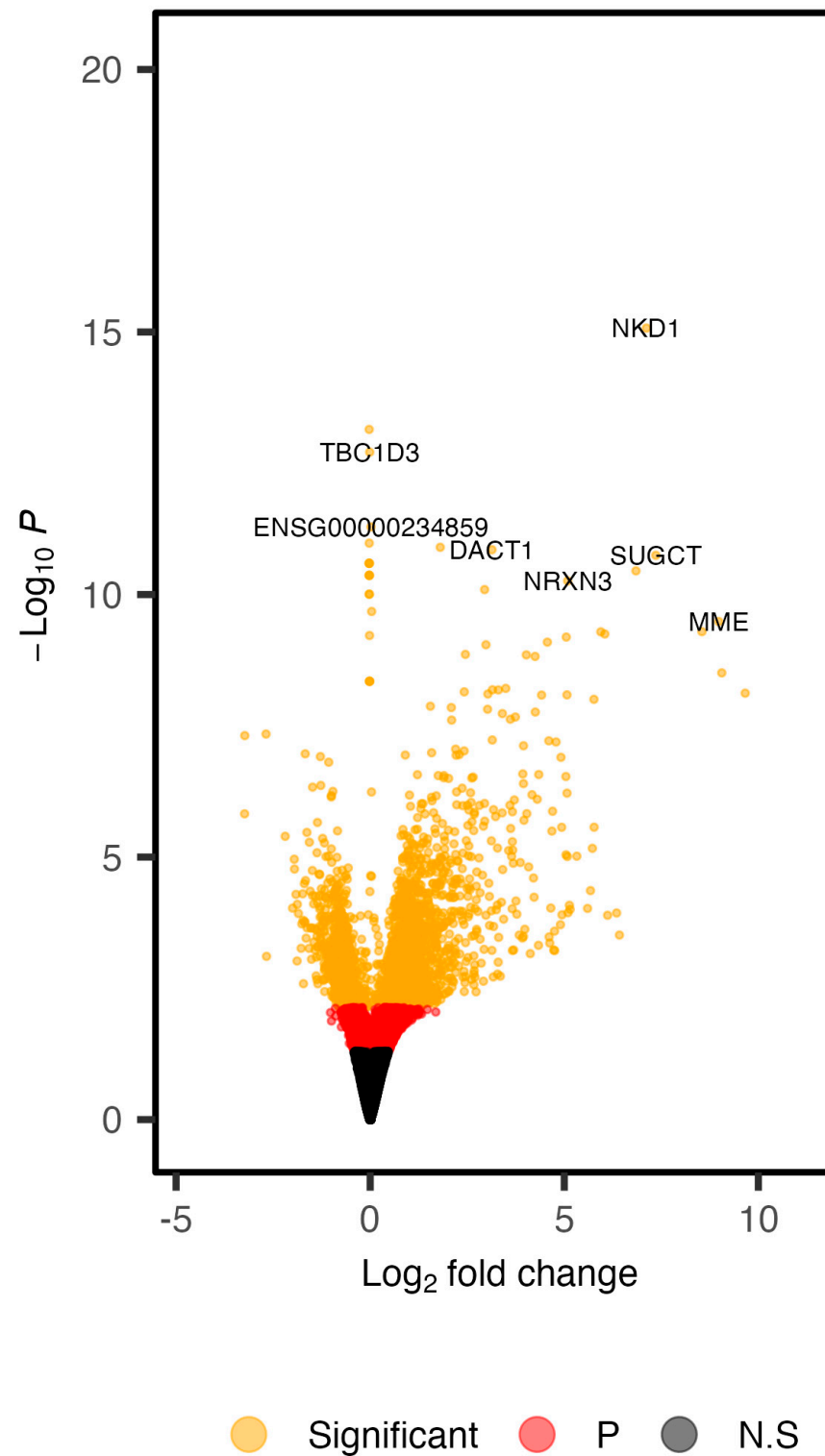
Figure 1. Summary of differential expression findings from our analysis of FAP versus healthy colon organoids. Genes were separated into three categories. Black genes (N.S) were those that did not reach any significance level. Red genes reached nominal significance, but did not survive multiple testing corrections. Generalized linear models were fit for each gene in DESeq2 and a Wald test was used for the identification of significant DEGs. 2635 DEGs (orange) genes remained significant following FDR correction (FDR < 0.05). Positive log2 fold changes represent genes that were increased in FAP versus healthy colon organoids.
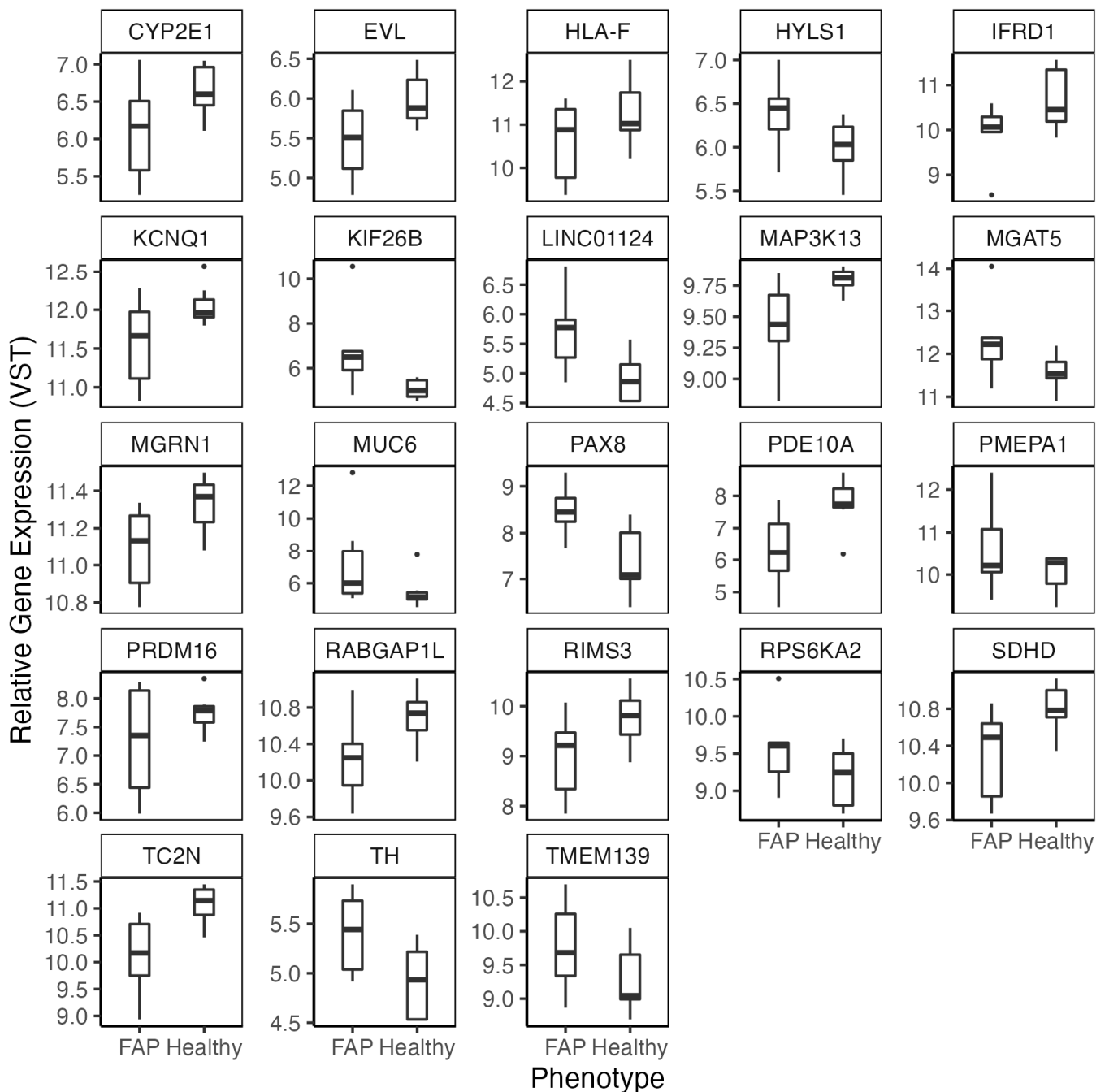
Figure 2. Boxplots of significant DEGs (FDR < 0.05) identified in our analysis (*n* = 13) that were previously found to be associated with the FAP phenotype through DNA methylation analysis of colon organoids. Generalized linear models were fit for each gene in DESeq2 and a Wald test was used for the identification of significant DEGs. To aid in visualization in the plotting genes, relative gene expression was calculated using the VST function in DESeq2.

3.3. Defining the Overlap in Tumor-Related Gene Expression between FAP and EOCRC

We next defined the transcriptomic landscapes of TCGA-EOCRC and FAP tumors using publicly available data. We performed paired regression analysis within each age grouping to identify DEGs between tumor versus NAT for each of the two subsets. A remarkable level of concordance was observed across the age groups. Indeed, homeobox B13 (*HOXB13*) was the only DEG found to be significant in both tumor versus NAT analyses that was discordant for direction of effect (Supplementary Figure S1 and Table S5).

To define the transcriptomic landscape of FAP tumors, we downloaded matched FAP tumor versus NAT data from GSE153385 and performed paired regression analyses in a manner similar to TCGA-COAD. We identified 951 of 2263 FAP tumor DEGs (42.01%) that were significant and concordant for direction of effect in our TCGA-EOCRC analysis, indicating a strong overlap for tumor related processes in the development of CRC for FAP and TCGA-EOCRC (Supplementary Table S6).

3.4. Overlap between DEGs Identified in FAP Colon Organoids and EOCRC Tumors

To determine the extent to which our FAP colon organoid analyses modeled the gene expression differences seen in EOCRC tumors, we overlaid our novel organoid findings with both FAP and TCGA-EOCRC tumors. Of the 2635 DEGs identified in our FAP colon organoid study, 210 were differentially expressed (FDR < 0.05) in FAP tumors versus NAT (Supplementary Table S7). This represented an enrichment of FAP tumor related DEGs in our analysis of colon organoids of FAP versus healthy subjects ($p = 1.22 \times 10^{-39}$). Next, we found that 38.41% (1012) of FAP colon organoid DEGs were present in at least one TCGA analysis (Supplementary Table S8). Enrichment for FAP colon organoid DEGs were observed in our analyses of TCGA-EOCRC ($p = 4.84 \times 10^{-8}$) and TCGA-60+ ($p = 7.17 \times 10^{-16}$). A total of nine DEGs unique to TCGA-EOCRC were identified in our FAP colon organoid analysis. Of these nine, only *EPDR1* was also present in FAP tumors (FDR = 0.01). However, *EPDR1* (FDR = 0.063) and forkhead box O3 (*FOXO3*; FDR = 0.051) and adrenomedullin (*ADM*; FDR = 0.056) trended towards significance in TCGA-60+. Finally, by comparing data generated across all analyses, we found that 142 DEGs were significant across each analysis (Figure 3; Supplementary Table S9). The majority of these DEGs (66.20%) were concordant for direction of effect in each dataset, with *FOXQ1* being the highest cumulatively ranked DEG observed across all four analyses.

3.5. Technical Validation of FAP DEGs and Confirmation of Findings in Biopsies

We performed qPCR on a subset of genes that were either among the most significant DEGs identified in our analysis of colon organoids of FAP versus healthy subjects and/or were significant in an analysis of either EOCRC tumor datasets: *NKD1, EYA2, EGR2, EPDR1* and *IGFL1*. Of the five genes chosen for technical validation, four were significant and followed the same direction of effect in our qPCR analysis (Table 1), while *IGFL1* displayed a trend towards a significant increase in FAP colon organoids ($p = 0.052$). By extending our analysis to matched FAP ($n = 2$) and healthy control ($n = 2$) colon biopsies, we found that three of the four genes (*EPDR1, EYA2* and *NKD1*) with detectable levels of expression were concordant for direction of effect (Supplementary Figure S2). Of note, *IGFL1* expression was not detected in FAP or healthy colon biopsies.

Table 1. Summary of qPCR findings for selected genes chosen for technical validation. CT values were generated for each gene following subtraction from internal control. High CT values correspond to low expression. Negative effect size estimates and test statistics therefore indicate increased expression of a gene in FAP versus healthy colon organoids.

| Gene	Estimate	Standard Error	Test Statistic	p
NKD1	−6.069	1.012	−6.001	5.42×10^{-4}
EGR2	−3.139	0.653	−4.805	1.95×10^{-3}
EYA2	3.820	1.326	2.880	0.024
EPDR1	−2.054	0.823	−2.496	0.047
IGFL1	−2.697	1.116	−2.417	0.052
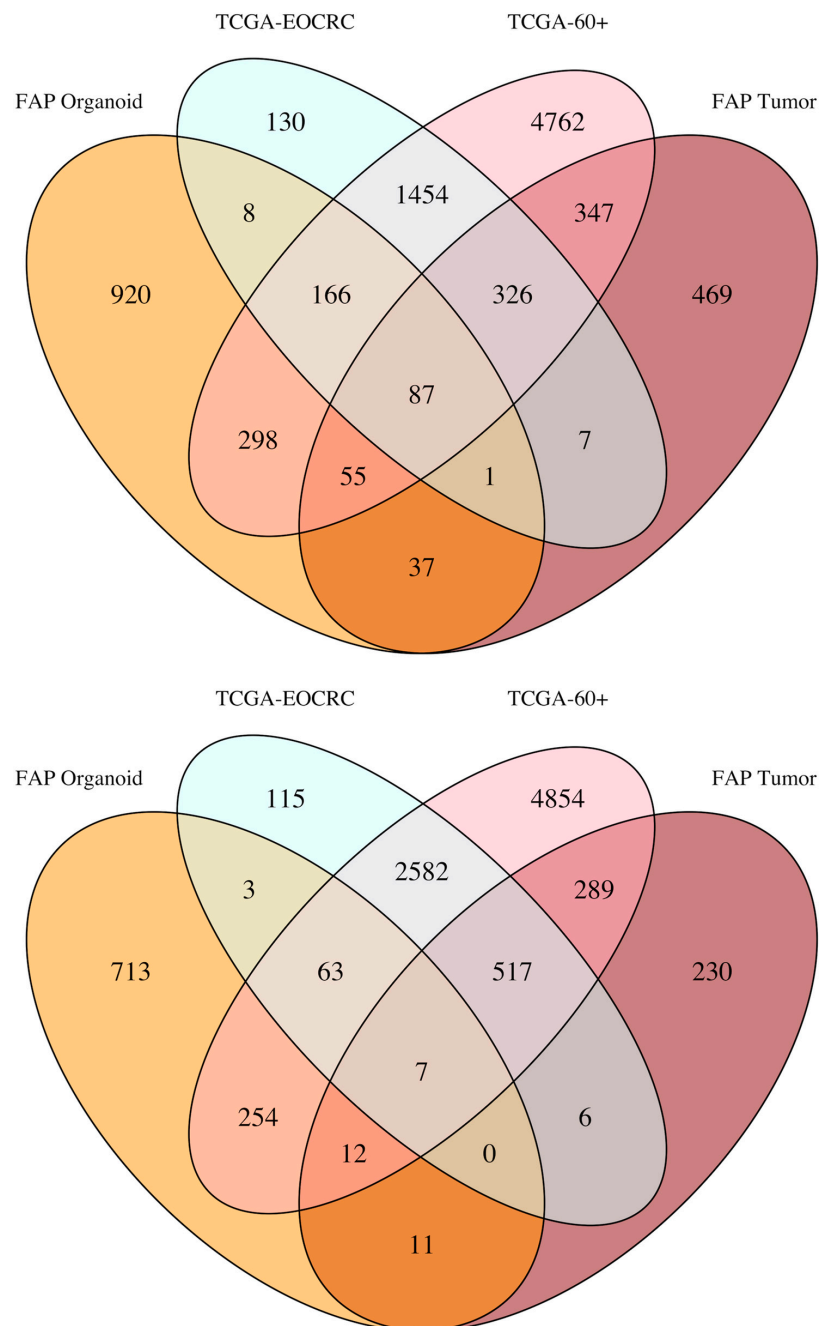
Figure 3. Venn diagram demonstrating the overlap of significant DEGs (FDR < 0.05) that were increased (left) and decreased (right) in FAP and CRC tumors in the four cohorts analyzed. Datasets were grouped by color: orange (FAP versus healthy colon organoids), light cyan (TCGA-EOCRC), pink (TCGA-60+) and dark red (FAP colon tumors (GSE153385)). Generalized linear models were fit for each gene in DESeq2 and a Wald test was used for the identification of significant DEGs.

3.6. Role of Lifestyle Factors in Driving Expression of FAP DEGs

A number of studies have implicated generational shifts in lifestyle factors as important contributors to the rise of EOCRC within Western populations [10]. We have previously shown that exposure of normal colon organoids to several different CRC related lifestyle factors led to widespread gene expression differences that are enriched for genes associated with CRC [18–22]. DEGs identified between organoids of FAP and healthy subjects were enriched for DEGs found to be differential expressed following ethanol ($p = 1.18 \times 10^{-62}$), aspirin ($p = 2.07 \times 10^{-19}$), calcium ($p = 2.36 \times 10^{-32}$) or carcinogen

cocktail ($p = 2.33 \times 10^{-77}$) exposure. Importantly, several of these genes were also identified in at least one of our EOCRC analysis. High levels of alcohol consumption have previously been linked to the development of high-risk adenomas and CRC in patients with at least one existing colorectal adenoma [43]. We therefore asked if FAP colon organoids exposed to ethanol led to further changes in expression of a subset of these genes in FAP organoids. Genes selected for targeted analysis were those previously shown to be differentially expressed in ethanol exposure (*NOTUM*, *DRAM1*, *PLEKHB1* and *GDF11*) [20] and in our analysis of FAP versus healthy colon organoids. Increased expression of *NOTUM* (FDR = 9.81×10^{-14}), *PLEKHB1* (FDR = 1.27×10^{-3}), *DRAM1* (FDR = 3.58×10^{-3}), *GDF11* (FDR = 0.012) were also identified in TCGA-EOCRC. Ethanol treatment led to increased expression of *NOTUM* in five of six, and increased expression of *GDF11* in all six FAP colon organoid lines (Figure 4), though effect sizes did vary. These increases were consistent for direction of effect in the larger, previous RNA-seq analysis of ethanol in healthy colon organoids [20]. More variable effects were observed for *PLEKHB1* and *DRAM1*, and larger sample sizes may be needed for the study of these genes.
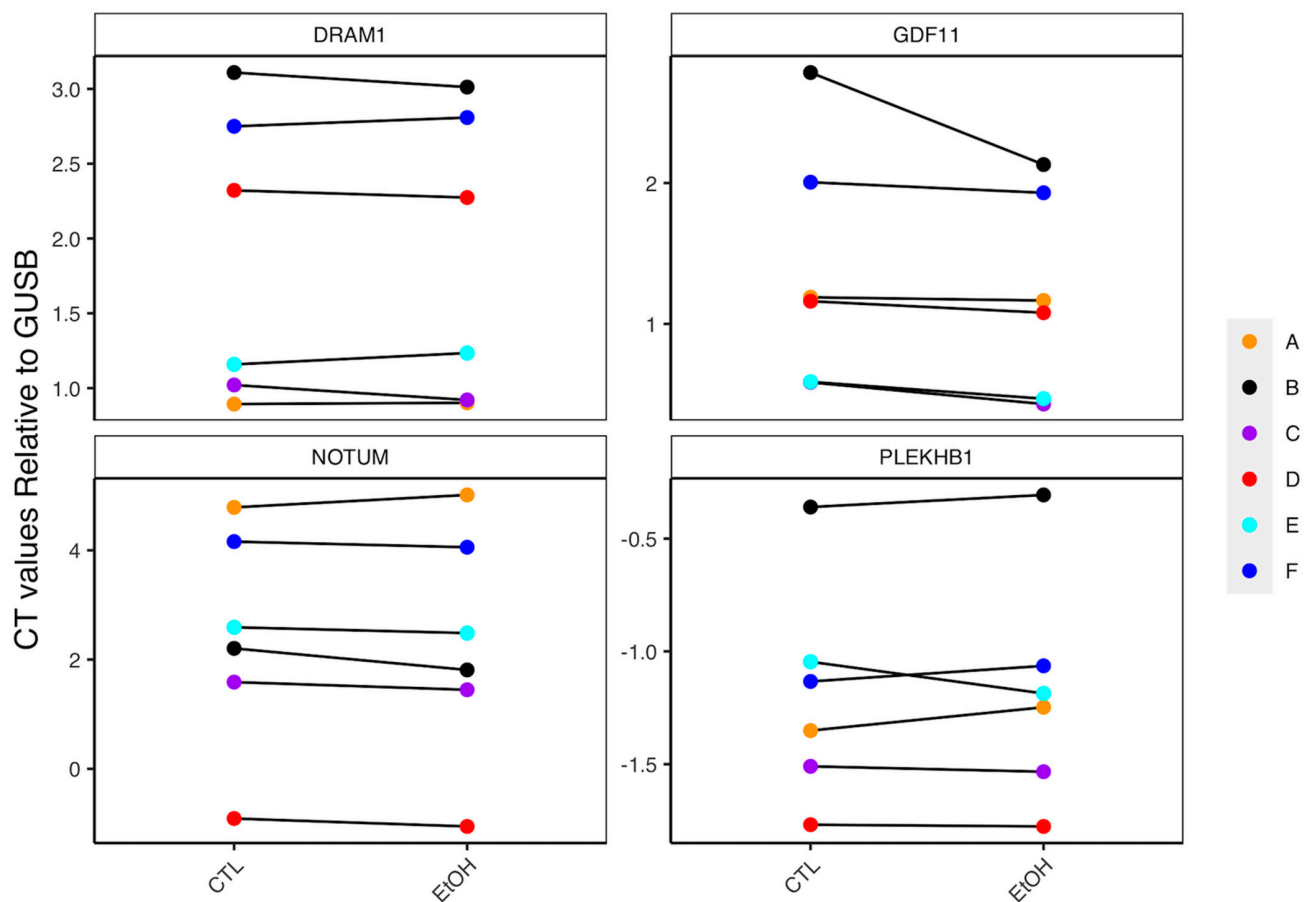
Figure 4. Summary of qPCR analysis from ethanol exposure of FAP colon organoid lines. CT values were generated for each gene following subtraction from internal control using a delta CT method. High CT values correspond to low expression. Reductions in CT scores between control (CTL) and ethanol treated (EtOH) FAP samples therefore represent an increase in expression for that gene in ethanol treated colon organoids. Paired samples (A–F) across treatment conditions are connected by a line to demonstrate overall direction of effect. Given the small sample size of the analysis, consistency for direction of effect was considered over significance testing to avoid the possibility that differential expression may be driven by subject-specific outliers.

4. Discussion

We have used the hereditary CRC syndrome, FAP, in an attempt to provide insight into molecular mechanisms involved in EOCRC development. We observed significant enrichments for DEGs identified in our analysis of colon organoids of FAP versus healthy subjects with those seen in FAP CRC tumors versus NAT (as expected), TCGA-EOCRC and TCGA-60+, which support our use of the FAP colon organoid system to model early events preceding EOCRC development. Additional functional work will be needed to confirm whether expression of these genes drive oncogenic signaling pathways.

We identified DEGs between colon organoids of FAP and healthy subjects that also differed consistently in an independent analysis of FAP tumor versus NAT. For example, a significant reduction in maternally expressed 3 (*MEG3*) expression was observed in both colon organoids of FAP patients and in FAP tumors. While this gene only showed a nominal decrease in TCGA-EOCRC tumors, no significant difference was observed in our analysis of TCGA-60+. Studies using CRC cancer cell lines have suggested that *MEG3* plays an important role as a tumor suppressor gene [44]. A recent study in *Apc^min* mice of colon stem cells suggested that *MEG3* inhibits early CRC development by acting as a sponge to *miR-708* [45]. While the authors conclude that further studies modifying *MEG3* expression in CRC tumors should be considered, our data suggests that studies prioritizing the modulation of *MEG3* expression in EOCRC tumors should also be considered.

To explore potential mechanisms underlying differential expression observed within our study, we aimed to determine the extent by which the DEGs identified here overlaid DMRs identified in our previously published, overlapping cohort of FAP and healthy colon organoids [42]. Aberrant levels of DNA methylation are considered to be a hallmark of CRC and many studies have identified widespread differential methylation in CRC tumors [16,39,46,47]. During CRC development, two seemingly contrary events unfold as extensive DNA hypomethylation and site-specific hypermethylation of DNA promoter regions lead to CRC-related events that contribute to the disease [48,49] Such differences have also led to aberrant expression of CRC-relevant genes [50] Here, we identified a subset of DEGs (*n* = 23) corresponding to previously DMRs that were also found to be specific to EOCRC [39]. This included the tumor suppressor gene, succinate dehydrogenase complex subunit D (*SDHD*), expression of which was found to be reduced in FAP colon organoids. However, while we identified differential expression in FAP colon organoids, the effects of this differential expression were greater in TCGA-60+ than TCGA-EOCRC. Despite the unique finding of differential methylation corresponding to EOCRC in previous studies [39] it is unclear whether the mechanisms driving expression of *SDHD* differ in EOCRC than later-onset CRC. Of the 23 genes that were identified, only *HLA g* was found to be at least nominally significant in TCGA-EOCRC and FAP colon organoids, but not TCGA-60+. However, given the highly diverse landscape of *HLA* and the relatively small sample sizes involved in this study, larger studies should first aim to replicate this finding.

FAP cases are rare, but we attempt to link data generated in colon organoids of FAP patients with biopsy data generated from a small subset of matched individuals. The colon organoid model displays greatly reduced cellular heterogeneity versus the biopsies from which they were derived, given that they lack immune and stromal cell populations. This is important given the spatiotemporal nature of gene expression. Indeed, the colon organoid model is representative of the stem cell niche of the colon crypt, which is considered to be the site of CRC development [51]. Thus, performing an analysis of the specific cells implicated in CRC tumor development provides relevance and reduces the risk of confounding in downstream regression analysis that may accompany increased cellular heterogeneity. However, the time, infrastructure and specific skill set required to generate organoids from colon biopsies limit their applicability as a model for biomarker evaluation. As such, we aimed to validate five DEGs identified through our colon organoid analysis in a small subset of colon biopsies. Three of the four genes tested were concordant for direction of effect with our colon organoid analyses. This included an increase in the expression of *EPDR1* in FAP colon organoids. Increased *EPDR1* expression was most

significant in EOCRC, reaching FDR correction in both FAP and TCGA-EOCRC analyses. However, this gene did trend to significance in our TCGA-60+ dataset and may therefore represent a strong target for early monitoring of CRC across age groups. That these DEGs were seen in both the colon organoid model and in a subset of colon biopsies of matched individuals highlights their relevance to CRC and the biomarker discovery potential of our study. Of interest, an increase in *IGFL1* was identified in our RNA-seq analysis of FAP colon organoids. *IGFL1* has previously been proposed to be an oncogene in other cancers, where it was found to aid in cancer cell growth and survival [52]. This increase was not found within TCGA-60+, but was identified in our TCGA-EOCRC analysis. Importantly, qPCR analysis of FAP colon biopsies was unable to detect expression of *IGFL1*. However, we were able to provide a partial technical validation for this finding in qPCR analysis of colon organoids of FAP versus healthy subjects. Tumor samples are known to have increased stemness scores versus matched biopsies [53], whereas colon organoids represent an important model of the colon stem cell niche. It is therefore possible that *IGFL1* plays an important role in stem cells and/or other undifferentiated cell populations that are generally of low abundance in normal biopsies. Indeed, previous studies have found that *IGFL1* expression is primarily limited within normal tissues that contain a high abundance of undifferentiated cells, such as fetal tissue and the spinal cord [54]. Importantly, IGFL sequences are not well conserved between mice and human [55], limiting its potential for modelling in other species. The identification of increased *IGFL1* expression in FAP colon organoids therefore supports the use of the organoid model for studying EOCRC. Future research would benefit from the availability of larger, publicly available EOCRC datasets for comparison. These datasets could be used to help better define DEGs unique to EOCRC development.

Previously, we have shown that exposure of CRC lifestyle factors leads to modified gene expression of CRC-related genes [18–22]. Given the strong enrichment of the FAP DEGs observed here and CRC, we were not surprised to see that FAP DEGs were also enriched for targets of various CRC lifestyle factors. Importantly, one theory attempting to explain the rise in EOCRC rates observed across recent generations centers on changes in Western lifestyle, including increased alcohol consumption [10]. As such, we performed qPCR analysis on a subset of genes previously found to be altered following ethanol treatment that were also differentially expressed in FAP colon organoids and in at least one EOCRC dataset. We identified a relatively consistent increase in the expression of *NOTUM*, as well as a consistent increase in the expression of *GDF11*. These findings were in line with the direction of effect observed in a larger analysis of normal colon organoids and represent a subset of EOCRC relevant DEGs whose expression are further perturbed following exposure to at least one common Western lifestyle factor. *GDF11* has also previously been associated with aging in numerous studies; however, its exact pathophysiological role remains unclear [56]. Increased *GDF11* has been associated with increased stemness [57], and CRC patients with increased expression of this genes have been shown to have reduced survival [58]. Silencing *GDF11* in the CRC cell line Caco-2 has also implicated a role for this gene in proliferation [57,59]. These findings add further weight to the potential role of known lifestyle factors in EOCRC. Further lifestyle factor exposure studies in organoids may yield improved insight into their role in both FAP EOCRC tumor development.

There are several limitations to our study. The first is the sample size of the datasets considered here. FAP sample size considerations were largely limited by the rarity of this syndrome in the population. Future studies should aim to build upon these findings through the use of larger cohorts. Second, we identify established modifiable targets of CRC relevant lifestyle factors through a limited selection process of genes that were also found to be related to CRC. It is unknown whether long term culture under these exposures may affect these genes further, or whether other genes that are relevant to CRC may be better identified. Further research is needed to explore how these genes may contribute to CRC risk. Third, we identify numerous genes relevant to FAP and TCGA-EOCRC tumorigenesis that are also seen in colon organoids of FAP patients in the absence of cancer. However, it

remains to be determined whether these DEGs reflect useful early biomarkers for EOCRC tumorigenesis and whether they have actionable potential as future drug targets. Further research will be needed in relevant models of CRC to determine whether these genes actively contribute to tumorigenesis. Determining whether these transcriptomic events are also present in blood would constitute an approach that would be more readily accessible and less invasive than biopsies collected at colonoscopy. Fourth, the normal-appearing FAP mucosal biopsies from which colon organoids were derived may have contained some, undetected microadenomas. Finally, inconsistent evidence for the role of APC in non-hereditary EOCRC exists. For example, the rate of APC mutations at known hotspot regions has been shown to be greater in non-hereditary, non-EOCRC populations [60]. However, sequencing of the *APC* gene in a small cohort of EOCRC patients found that mutation rates across the gene occurred at a similar level as proposed in non-hereditary, non-EOCRC populations. Most importantly, the authors conclude that the pattern of these mutations were more similar to inherited *APC* mutations than to non-hereditary, later onset CRC [60]

In conclusion, we have identified widespread differential expression in colon organoids of FAP and healthy subjects. These DEGs were enriched for genes related to EOCRC. Our findings highlight potential early markers for EOCRC, but also mechanisms through which lifestyle factors may further modulate the expression of genes implicated in EOCRC.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/cancers14174138/s1, Supplementary Table S1: Relevant demographic information of the colon organoids used within this study. Supplementary Table S2: RNA-seq mapping statistics for all novel samples used in the study. Supplementary Table S3: Pathway enrichment analysis of significant DEGs identified in regression of FAP versus healthy colon organoids. Supplementary Table S4: Summary of RNA-seq findings for a subset of genes for which DMRs had previously been identified. Positive fold changes represent genes that were increased in FAP versus healthy colon organoids. Supplementary Table S5: Summary of RNA-seq findings for analysis of non-hereditary EOCRC and non-EOCRC regression models. Positive fold changes represent genes that were increased in tumor versus NAT. Cumulative rank was generated by first ranking significant DEGs in each analysis in order of greatest significance and summing the total of these two values. A rank was then ordered based on the subsequent lowest scores. Supplementary Table S6: Summary of consistent RNA-seq findings identified in FAP tumor and non-hereditary EOCRC regression models. Positive fold changes represent genes that were increased in tumor versus NAT. Cumulative rank was generated by first ranking significant DEGs in each analysis in order of greatest significance and summing the total of these two values. A rank was then ordered based on the subsequent lowest scores. Only DEGs that were significant in both analyses were reported. Supplementary Table S7: Summary of RNA-seq consistent findings identified in FAP colon organoid and FAP colon tumor regression models. Positive fold changes represent genes that were increased in FAP versus healthy colon organoids and/or tumor versus NAT. Cumulative rank was generated by first ranking significant DEGs in each analysis in order of greatest significance and summing the total of these two values. A rank was then ordered based on the subsequent lowest scores. Only DEGs that were significant in both analyses were reported. Supplementary Table S8: Summary of RNA-seq findings identified in FAP colon organoids and both non-hereditary EOCRC regression models. Positive fold changes represent genes that were increased in FAP versus healthy colon organoids and/or tumor versus NAT. Cumulative rank was generated by first ranking significant DEGs in each analysis in order of greatest significance and summing the total of these three values. A rank was then ordered based on the subsequent lowest scores. Supplementary Table S9: Summary of significant DEGs identified from RNA-seq in each analysis. Positive fold changes represent genes that were increased in FAP versus healthy colon organoids and/or tumor versus NAT. Cumulative rank was generated by first ranking significant DEGs in each analysis in order of greatest significance and summing the total of these three values. A rank was then ordered based on the subsequent lowest scores. Only genes that were concordant for direction of effect in each analysis were ranked. Supplementary Figure S1: Volcano plot of DEGs identified in FAP versus healthy colon organoid analysis (*n* = 13) that demonstrates highlights similarity of differential expression reporting in EOCRC tumor datasets. Genes were separated into five categories. Black genes (N.S) were those that did not reach any

significance level. Orange genes reached FDR significance, but were unique to FAP colon organoids. Dark red genes were significant in colon organoids and TCGA-EOCRC. Purple genes were significant in colon organoids and FAP tumors. Blue genes were significant in colon organoids, TCGA-EOCRC and FAP tumors. Generalized linear models were fit for each gene in DESeq2 and a Wald test was used for the identification of significant DEGs. Positive log2 fold changes represent genes that were increased in FAP versus healthy colon organoids. Supplementary Figure S2: QPCR analysis of FAP (*n* = 2) versus healthy colon biopsies (*n* = 2). Following normalization to internal control, the relative expression of each gene, for each FAP patient was calculated against the average expression of both healthy controls. Values greater than one indicated increased expression in FAP colon biopsies. *IGFL1* expression was not detected in any biopsy.

Author Contributions: Conceptualization, M.A.D. and G.C.; Methodology, M.A.D., G.C., M.W.A. and S.E.; Software, M.A.D.; Validation, M.A.D., S.E. and M.W.A.; Formal Analysis, M.A.D.; Investigation, all authors; Resources, M.A.D., S.E., M.W.A., S.M.P. and G.C.; Data Curation, S.E., S.M.P. and M.A.D.; Writing—Original Draft Preparation, M.A.D., G.C. and L.L.; Writing—Review and Editing, all authors; Visualization, M.A.D.; Supervision, G.C.; Project Administration, G.C.; Funding Acquisition, G.C. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Review Board of University of Virginia (IRB-HSR #19439).

Informed Consent Statement: Written informed consent has been obtained from the patient(s) to publish this paper and the study was conducted in accordance with U.S. Common Rule.

Data Availability Statement: All data has been made available to Gene Expression Omnibus and can be accessed through accession number: GSE207398.

Conflicts of Interest: The authors declare no conflict of interest. The sponsors had no role in the design, execution, interpretation, or writing of the study.

References

1. Huyghe, J.R.; Bien, S.A.; Harrison, T.A.; Kang, H.M.; Chen, S.; Schmit, S.L.; Conti, D.V.; Qu, C.; Jeon, J.; Edlund, C.K.; et al. Discovery of common and rare genetic risk variants for colorectal cancer. *Nat. Genet.* **2019**, *51*, 76–87. [CrossRef] [PubMed]

2. Schmit, S.L.; Edlund, C.K.; Schumacher, F.R.; Gong, J.; Harrison, T.A.; Huyghe, J.R.; Qu, C.; Melas, M.; Van Den Berg, D.J.; Wang, H.; et al. Novel Common Genetic Susceptibility Loci for Colorectal Cancer. *J. Natl. Cancer Inst.* **2019**, *111*, 146–157. [CrossRef] [PubMed]

3. Law, P.J.; Timofeeva, M.; Fernandez-Rozadilla, C.; Broderick, P.; Studd, J.; Fernandez-Tajes, J.; Farrington, S.; Svinti, V.; Palles, C.; Orlando, G.; et al. Association analyses identify 31 new risk loci for colorectal cancer susceptibility. *Nat. Commun.* **2019**, *10*, 2154. [CrossRef] [PubMed]

4. 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] [PubMed]

5. Bailey, C.E.; Hu, C.Y.; You, Y.N.; Bednarski, B.K.; Rodriguez-Bigas, M.A.; Skibber, J.M.; Cantor, S.B.; Chang, G.J. Increasing disparities in the age-related incidences of colon and rectal cancers in the United States, 1975-2010. *JAMA Surg.* **2015**, *150*, 17–22. [CrossRef] [PubMed]

6. 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]

7. Force, U.S.P.S.T.; Davidson, K.W.; Barry, M.J.; Mangione, C.M.; Cabana, M.; Caughey, A.B.; Davis, E.M.; Donahue, K.E.; Doubeni, C.A.; Krist, A.H.; et al. Screening for Colorectal Cancer: US Preventive Services Task Force Recommendation Statement. *JAMA* **2021**, *325*, 1965–1977. [CrossRef]

8. Vuik, F.E.; Nieuwenburg, S.A.; Bardou, M.; Lansdorp-Vogelaar, I.; Dinis-Ribeiro, M.; Bento, M.J.; Zadnik, V.; Pellise, M.; Esteban, L.; Kaminski, M.F.; et al. Increasing incidence of colorectal cancer in young adults in Europe over the last 25 years. *Gut* **2019**, *68*, 1820–1826. [CrossRef]

9. Patel, S.G.; Karlitz, J.J.; Yen, T.; Lieu, C.H.; Boland, C.R. The rising tide of early-onset colorectal cancer: A comprehensive review of epidemiology, clinical features, biology, risk factors, prevention, and early detection. *Lancet Gastroenterol. Hepatol.* **2022**, *7*, 262–274. [CrossRef]

10. Wu, C.W.; Lui, R.N. Early-onset colorectal cancer: Current insights and future directions. *World J. Gastrointest. Oncol.* **2022**, *14*, 230–241. [CrossRef]

11. Chen, X.; Li, H.; Guo, F.; Hoffmeister, M.; Brenner, H. Alcohol consumption, polygenic risk score, and early- and late-onset colorectal cancer risk. *eClinicalMedicine* **2022**, *49*, 101460. [CrossRef] [PubMed]

12. Stec, R.; Plawski, A.; Synowiec, A.; Maczewski, M.; Szczylik, C. Colorectal cancer in the course of familial adenomatous polyposis syndrome ("de novo" pathogenic mutation of APC gene): Case report, review of the literature and genetic commentary. *Arch. Med. Sci.* **2010**, *6*, 283–287. [CrossRef] [PubMed]

13. Groden, J.; Thliveris, A.; Samowitz, W.; Carlson, M.; Gelbert, L.; Albertsen, H.; Joslyn, G.; Stevens, J.; Spirio, L.; Robertson, M.; et al. Identification and characterization of the familial adenomatous polyposis coli gene. *Cell* **1991**, *66*, 589–600. [CrossRef]

14. Stoffel, E.M.; Koeppe, E.; Everett, J.; Ulintz, P.; Kiel, M.; Osborne, J.; Williams, L.; Hanson, K.; Gruber, S.B.; Rozek, L.S. Germline Genetic Features of Young Individuals With Colorectal Cancer. *Gastroenterology* **2018**, *154*, 897–905.e1. [CrossRef] [PubMed]

15. Mork, M.E.; You, Y.N.; Ying, J.; Bannon, S.A.; Lynch, P.M.; Rodriguez-Bigas, M.A.; Vilar, E. High Prevalence of Hereditary Cancer Syndromes in Adolescents and Young Adults With Colorectal Cancer. *J. Clin. Oncol.* **2015**, *33*, 3544–3549. [CrossRef] [PubMed]

16. Cancer Genome Atlas, N. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* **2012**, *487*, 330–337. [CrossRef]

17. Barker, N.; Ridgway, R.A.; van Es, J.H.; van de Wetering, M.; Begthel, H.; van den Born, M.; Danenberg, E.; Clarke, A.R.; Sansom, O.J.; Clevers, H. Crypt stem cells as the cells-of-origin of intestinal cancer. *Nature* **2009**, *457*, 608–611. [CrossRef]

18. Devall, M.; Dampier, C.H.; Eaton, S.; Ali, M.W.; Díez-Obrero, V.; Moratalla-Navarro, F.; Bryant, J.; Jennelle, L.T.; Moreno, V.; Powell, S.M.; et al. Novel insights into the molecular mechanisms underlying risk of colorectal cancer from smoking and red/processed meat carcinogens by modeling exposure in normal colon organoids. *Oncotarget* **2021**, *12*, 1863. [CrossRef]

19. Devall, M.; Jennelle, L.T.; Bryant, J.; Bien, S.; Peters, U.; Powell, S.; Casey, G. Modeling the effect of prolonged ethanol exposure on global gene expression and chromatin accessibility in normal 3D colon organoids. *PLoS ONE* **2020**, *15*, e0227116. [CrossRef]

20. Devall, M.; Plummer, S.J.; Bryant, J.; Jennelle, L.T.; Eaton, S.; Dampier, C.H.; Huyghe, J.R.; Peters, U.; Powell, S.M.; Casey, G. Ethanol exposure drives colon location specific cell composition changes in a normal colon crypt 3D organoid model. *Sci. Rep.* **2021**, *11*, 432. [CrossRef]

21. Devall, M.A.M.; Drew, D.A.; Dampier, C.H.; Plummer, S.J.; Eaton, S.; Bryant, J.; Diez-Obrero, V.; Mo, J.; Kedrin, D.; Zerjav, D.C.; et al. Transcriptome-wide In Vitro Effects of Aspirin on Patient-derived Normal Colon Organoids. *Cancer Prev. Res.* **2021**, *14*, 1089–1100. [CrossRef]

22. Devall, A.M.; Dampier, C.H.; Eaton, S.; Ali, M.W.; Plummer, S.J.; Bryant, J.; Gauderman, W.J.; Peters, U.; Powell, S.M.; Casey, G. Transcriptomic response to calcium in normal colon organoids is impacted by colon location and sex. *Cancer Prev. Res.* **2022**, *9*, 2103360.

23. Frankish, A.; Diekhans, M.; Ferreira, A.M.; Johnson, R.; Jungreis, I.; Loveland, J.; Mudge, J.M.; Sisu, C.; Wright, J.; Armstrong, J.; et al. GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res.* **2019**, *47*, D766–D773. [CrossRef] [PubMed]

24. 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]

25. Li, B.; Dewey, C.N. RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinform.* **2011**, *12*, 323. [CrossRef]

26. Edgar, R.; Domrachev, M.; Lash, A.E. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* **2002**, *30*, 207–210. [CrossRef]

27. Bommi, P.V.; Bowen, C.M.; Reyes-Uribe, L.; Wu, W.; Katayama, H.; Rocha, P.; Parra, E.R.; Francisco-Cruz, A.; Ozcan, Z.; Tosti, E.; et al. The Transcriptomic Landscape of Mismatch Repair-Deficient Intestinal Stem Cells. *Cancer Res.* **2021**, *81*, 2760–2773. [CrossRef]

28. 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]

29. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2021.

30. Soneson, C.; Love, M.I.; Robinson, M.D. Differential analyses for RNA-seq: Transcript-level estimates improve gene-level inferences. *F1000Research* **2015**, *4*, 1521. [CrossRef] [PubMed]

31. 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]

32. Szklarczyk, D.; Gable, A.L.; Lyon, D.; Junge, A.; Wyder, S.; Huerta-Cepas, J.; Simonovic, M.; Doncheva, N.T.; Morris, J.H.; Bork, P.; et al. STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **2019**, *47*, D607–D613. [CrossRef]

33. Gene Ontology, C. The Gene Ontology resource: Enriching a GOld mine. *Nucleic Acids Res.* **2021**, *49*, D325–D334. [CrossRef] [PubMed]

34. Kanehisa, M.; Furumichi, M.; Sato, Y.; Ishiguro-Watanabe, M.; Tanabe, M. KEGG: Integrating viruses and cellular organisms. *Nucleic Acids Res.* **2021**, *49*, D545–D551. [CrossRef] [PubMed]

35. Zhu, A.; Ibrahim, J.G.; Love, M.I. Heavy-tailed prior distributions for sequence count data: Removing the noise and preserving large differences. *Bioinformatics* **2019**, *35*, 2084–2092. [CrossRef]

36. Blighe, K.; Rana, S.; Lewis, M. *EnhancedVolcano: Publication-Ready Volcano Plots with Enhanced Colouring and Labeling, 1.14.0*; R Core Team: Vienna, Austria, 2022.

37. Forsyth, C.B.; Tang, Y.; Shaikh, M.; Zhang, L.; Keshavarzian, A. Alcohol stimulates activation of Snail, epidermal growth factor receptor signaling, and biomarkers of epithelial-mesenchymal transition in colon and breast cancer cells. *Alcohol. Clin. Exp. Res.* **2010**, *34*, 19–31. [CrossRef]

38. Swanson, G.; Forsyth, C.B.; Tang, Y.; Shaikh, M.; Zhang, L.; Turek, F.W.; Keshavarzian, A. Role of intestinal circadian genes in alcohol-induced gut leakiness. *Alcohol. Clin. Exp. Res.* **2011**, *35*, 1305–1314. [CrossRef]

39. Joo, J.E.; Clendenning, M.; Wong, E.M.; Rosty, C.; Mahmood, K.; Georgeson, P.; Winship, I.M.; Preston, S.G.; Win, A.K.; Dugue, P.A.; et al. DNA Methylation Signatures and the Contribution of Age-Associated Methylomic Drift to Carcinogenesis in Early-Onset Colorectal Cancer. *Cancers* **2021**, *13*, 2589. [CrossRef]

40. Grant, A.; Xicola, R.M.; Nguyen, V.; Lim, J.; Thorne, C.; Salhia, B.; Llor, X.; Ellis, N.; Padi, M. Molecular drivers of tumor progression in microsatellite stable APC mutation-negative colorectal cancers. *Sci. Rep.* **2021**, *11*, 23507. [CrossRef]

41. Willauer, A.N.; Liu, Y.; Pereira, A.A.L.; Lam, M.; Morris, J.S.; Raghav, K.P.S.; Morris, V.K.; Menter, D.; Broaddus, R.; Meric-Bernstam, F.; et al. Clinical and molecular characterization of early-onset colorectal cancer. *Cancer* **2019**, *125*, 2002–2010. [CrossRef]

42. Devall, M.A.; Eaton, S.; Ali, M.W.; Dampier, C.H.; Weisenberger, D.J.; Powell, S.M.; Li, L.; Casey, G. DNA methylation analysis of normal colon organoids from familial adenomatous polyposis patients reveals novel insight into colon cancer development. *Clin. Epigenetics* **2022**, *14*, 104. [CrossRef]

43. Bardou, M.; Montembault, S.; Giraud, V.; Balian, A.; Borotto, E.; Houdayer, C.; Capron, F.; Chaput, J.C.; Naveau, S. Excessive alcohol consumption favours high risk polyp or colorectal cancer occurrence among patients with adenomas: A case control study. *Gut* **2002**, *50*, 38–42. [CrossRef] [PubMed]

44. Zhou, Y.; Zhong, Y.; Wang, Y.; Zhang, X.; Batista, D.L.; Gejman, R.; Ansell, P.J.; Zhao, J.; Weng, C.; Klibanski, A. Activation of p53 by MEG3 non-coding RNA. *J. Biol. Chem.* **2007**, *282*, 24731–24742. [CrossRef] [PubMed]

45. Zhang, S.; Ji, W.W.; Wei, W.; Zhan, L.X.; Huang, X. Long noncoding RNA Meg3 sponges miR-708 to inhibit intestinal tumorigenesis via SOCS3-repressed cancer stem cells growth. *Cell Death Dis.* **2021**, *13*, 25. [CrossRef] [PubMed]

46. Fan, J.; Li, J.; Guo, S.; Tao, C.; Zhang, H.; Wang, W.; Zhang, Y.; Zhang, D.; Ding, S.; Zeng, C. Genome-wide DNA methylation profiles of low- and high-grade adenoma reveals potential biomarkers for early detection of colorectal carcinoma. *Clin. Epigenetics* **2020**, *12*, 56. [CrossRef]

47. Park, S.K.; Song, C.S.; Yang, H.J.; Jung, Y.S.; Choi, K.Y.; Koo, D.H.; Kim, K.E.; Jeong, K.U.; Kim, H.O.; Kim, H.; et al. Field Cancerization in Sporadic Colon Cancer. *Gut Liver* **2016**, *10*, 773–780. [CrossRef]

48. Hidaka, H.; Higashimoto, K.; Aoki, S.; Mishima, H.; Hayashida, C.; Maeda, T.; Koga, Y.; Yatsuki, H.; Joh, K.; Noshiro, H.; et al. Comprehensive methylation analysis of imprinting-associated differentially methylated regions in colorectal cancer. *Clin. Epigenetics* **2018**, *10*, 150. [CrossRef]

49. Grady, W.M.; Carethers, J.M. Genomic and epigenetic instability in colorectal cancer pathogenesis. *Gastroenterology* **2008**, *135*, 1079–1099. [CrossRef]

50. Kerachian, M.A.; Javadmanesh, A.; Azghandi, M.; Mojtabanezhad Shariatpanahi, A.; Yassi, M.; Shams Davodly, E.; Talebi, A.; Khadangi, F.; Soltani, G.; Hayatbakhsh, A.; et al. Crosstalk between DNA methylation and gene expression in colorectal cancer, a potential plasma biomarker for tracing this tumor. *Sci. Rep.* **2020**, *10*, 2813. [CrossRef]

51. Sato, T.; Stange, D.E.; Ferrante, M.; Vries, R.G.; Van Es, J.H.; Van den Brink, S.; Van Houdt, W.J.; Pronk, A.; Van Gorp, J.; Siersema, P.D.; et al. Long-term expansion of epithelial organoids from human colon, adenoma, adenocarcinoma, and Barrett's epithelium. *Gastroenterology* **2011**, *141*, 1762–1772. [CrossRef]

52. Wang, H.; Shi, Y.; Chen, C.H.; Wen, Y.; Zhou, Z.; Yang, C.; Sun, J.; Du, G.; Wu, J.; Mao, X.; et al. KLF5-induced lncRNA IGFL2-AS1 promotes basal-like breast cancer cell growth and survival by upregulating the expression of IGFL1. *Cancer Lett.* **2021**, *515*, 49–62. [CrossRef]

53. Malta, T.M.; Sokolov, A.; Gentles, A.J.; Burzykowski, T.; Poisson, L.; Weinstein, J.N.; Kaminska, B.; Huelsken, J.; Omberg, L.; Gevaert, O.; et al. Machine Learning Identifies Stemness Features Associated with Oncogenic Dedifferentiation. *Cell* **2018**, *173*, 338–354.e315. [CrossRef] [PubMed]

54. Emtage, P.; Vatta, P.; Arterburn, M.; Muller, M.W.; Park, E.; Boyle, B.; Hazell, S.; Polizotto, R.; Funk, W.D.; Tang, Y.T. IGFL: A secreted family with conserved cysteine residues and similarities to the IGF superfamily. *Genomics* **2006**, *88*, 513–520. [CrossRef] [PubMed]

55. Lobito, A.A.; Ramani, S.R.; Tom, I.; Bazan, J.F.; Luis, E.; Fairbrother, W.J.; Ouyang, W.; Gonzalez, L.C. Murine insulin growth factor-like (IGFL) and human IGFL1 proteins are induced in inflammatory skin conditions and bind to a novel tumor necrosis factor receptor family member, IGFLR1. *J. Biol. Chem.* **2011**, *286*, 18969–18981. [CrossRef]

56. Anon-Hidalgo, J.; Catalan, V.; Rodriguez, A.; Ramirez, B.; Silva, C.; Galofre, J.C.; Salvador, J.; Fruhbeck, G.; Gomez-Ambrosi, J. Circulating GDF11 levels are decreased with age but are unchanged with obesity and type 2 diabetes. *Aging* **2019**, *11*, 1733–1744. [CrossRef]

57. Simoni-Nieves, A.; Gerardo-Ramirez, M.; Pedraza-Vazquez, G.; Chavez-Rodriguez, L.; Bucio, L.; Souza, V.; Miranda-Labra, R.U.; Gomez-Quiroz, L.E.; Gutierrez-Ruiz, M.C. GDF11 Implications in Cancer Biology and Metabolism. Facts and Controversies. *Front. Oncol.* **2019**, *9*, 1039. [CrossRef]

58. Yokoe, T.; Ohmachi, T.; Inoue, H.; Mimori, K.; Tanaka, F.; Kusunoki, M.; Mori, M. Clinical significance of growth differentiation factor 11 in colorectal cancer. *Int. J. Oncol.* **2007**, *31*, 1097–1101. [PubMed]

59. Barresi, V.; Reggiani-Bonetti, L.; Di Gregorio, C.; De Leon, M.P.; Barresi, G. Lymphatic vessel density and its prognostic value in stage I colorectal carcinoma. *J. Clin. Pathol.* **2011**, *64*, 6–12. [CrossRef]

60. Aitchison, A.; Hakkaart, C.; Day, R.C.; Morrin, H.R.; Frizelle, F.A.; Keenan, J.I. APC Mutations Are Not Confined to Hotspot Regions in Early-Onset Colorectal Cancer. *Cancers* **2020**, *12*, 3829. [CrossRef]