

Article

Age-Stratified Transcriptomic Profiling Identifies CEMIP as a Candidate Biomarker in Early-Onset Colorectal Cancer and Reveals an Association with PTCH1-Related Hedgehog Signaling

Chung-Ying Lee ^{1,2,†}, Hsu-Jui Pan ^{3,†}, Yu-Cheng Lee ⁴, Hieu Duc Nguyen ⁵, Yi-Chun Ni ⁵, Ke Xin Yee ⁵, Man Thi Nguyen ³, Yung-Fu Wu ^{6,7,*} and Kuen-Haur Lee ^{3,5,8,9,10,*}

- ¹ Division of Gastroenterology and Hepatology, Department of Internal Medicine, Taipei Medical University Shuang Ho Hospital Ministry of Health and Welfare, New Taipei City 23561, Taiwan; 12310@s.tmu.edu.tw
- ² Gastroenterology and Hepatology, Taipei Medical University School of Medicine, Taipei 11031, Taiwan
- ³ Graduate Institute of Cancer Molecular Biology and Drug Discovery, College of Medical Science and Technology, Taipei Medical University-Shuangho Campus, New Taipei City 23561, Taiwan; panray8802069487@gmail.com (H.-J.P.); mannguyenthi2001@gmail.com (M.T.N.)
- ⁴ Graduate Institute of Medical Sciences, College of Medicine, Taipei Medical University, Taipei 11031, Taiwan; ycle0212@tmu.edu.tw
- ⁵ PhD Program for Cancer Molecular Biology and Drug Discovery, College of Medical Science and Technology, Taipei Medical University-Shuangho Campus, New Taipei City 235, Taiwan; hieudn206@gmail.com (H.D.N.); ckn1012@gmail.com (Y.-C.N.); d621112004@tmu.edu.tw (K.X.Y.)
- ⁶ Department of Medical Research, Tri-Service General Hospital, National Defense Medical University, Taipei 114, Taiwan
- ⁷ Graduate Institute of Medical Sciences, College of Medicine, National Defense Medical University, Taipei 114, Taiwan
- ⁸ TMU Research Center for Digestive Medicine, Taipei Medical University, Taipei 110, Taiwan
- ⁹ Cancer Center, Wan Fang Hospital, Taipei Medical University, Taipei 110, Taiwan
- ¹⁰ TMU Research Center of Cancer Translational Medicine, Taipei Medical University, Taipei 110, Taiwan
- * Correspondence: prince1109@mail.ndmctsgh.edu.tw (Y.-F.W.); khlee@tmu.edu.tw (K.-H.L.); Tel.: +886-2-8792-3100 (ext. 19635) (Y.-F.W.); +886-2-26972035 (ext. 119) (K.-H.L.)
- † These authors contributed equally to this work.

Abstract

Background/Objectives: Early-onset colorectal cancer (EOCRC), defined as colorectal cancer diagnosed before 50 years of age, is increasing worldwide and may exhibit molecular features distinct from late-onset colorectal cancer (LOCRC). This study aimed to identify EOCRC-associated genes and evaluate the functional relevance of the leading candidate. **Methods:** Public transcriptomic datasets were used for age-stratified candidate-gene discovery and validation. Overall survival analysis prioritized candidates, followed by loss-of-function studies and pathway-focused analyses in colorectal cancer cell lines. **Results:** CEMIP, NKD2, and FOXQ1 were elevated in the EOCRC groups in the discovery and integrated validation analyses. Among them, only CEMIP was significantly associated with poorer overall survival. HCT116 and HT29 cells showed relatively high endogenous CEMIP expression, and CEMIP knockdown reduced relative viable cell biomass and migration in wound-healing and Transwell assays. CEMIP-associated transcripts were enriched in several pathways, including Hedgehog, Wnt/ β -catenin, and TGF- β signaling. In TCGA colon adenocarcinoma samples, CEMIP correlated most strongly with PTCH1; however, the correlations with other Hedgehog components were weak or absent. CEMIP silencing decreased PTCH1 and GLI1 transcripts while increasing SMO, suggesting a co-

Academic Editors: Lucio Bertario and Xin Wang

Received: 20 June 2026

Revised: 30 July 2026

Accepted: 3 August 2026

Published: 4 August 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\) license](https://creativecommons.org/licenses/by/4.0/).

ordinated but non-linear relationship. Conclusions: CEMIP is an EOCRC-associated candidate biomarker and a functional contributor to colorectal cancer cell viability and migration. Its relationship with PTCH1-related Hedgehog signaling is exploratory and warrants direct mechanistic validation.

Keywords: early-onset colorectal cancer; CEMIP; Hedgehog signaling; transcriptomics

1. Introduction

Colorectal cancer (CRC) is the third most common malignancy and the second leading cause of cancer-related death worldwide, accounting for approximately 1.9 million new cases and 900,000 deaths in 2020 [1]. More than 90% of CRCs are adenocarcinomas arising from the glandular epithelium of the colon and rectum [2]. Most CRC cases are sporadic and develop through multistep tumorigenic processes involving adenomatous or serrated precursor lesions, including the adenoma-carcinoma sequence, the serrated pathway, and, less commonly, inflammation-associated tumorigenesis [3–8]. Although the incidence and mortality of CRC in older adults have gradually declined owing to screening, early detection, and treatment advances [9], the incidence of early-onset colorectal cancer (EOCRC), generally defined as CRC diagnosed before the age of 50 years, has continued to increase over the past two decades [10–15]. A large European cohort study further showed that CRC incidence increased annually by 7.9% in individuals in their 20s, 4.9% in those in their 30s, and 1.6% in those in their 40s, underscoring the growing clinical importance of EOCRC [16,17].

EOCRC is increasingly recognized as a distinct and heterogeneous disease entity. Although some cases are attributable to hereditary cancer syndromes or familial predisposition, more than 80% of patients with EOCRC do not fall into these categories [10,11]. Compared with late-onset colorectal cancer (LOCRC), EOCRC is more frequently diagnosed at an advanced stage and is often associated with more aggressive clinicopathological features [18,19]. These observations suggest that EOCRC may harbor distinct molecular characteristics; however, no unified molecular mechanism or specific biomarker has yet been established [9]. Because tumor metastasis is a major cause of cancer-related mortality, pathways involved in epithelial–mesenchymal transition (EMT), invasion, and dissemination are of particular interest in EOCRC [20]. EMT is regulated by several oncogenic pathways, including TGF- β , Wnt, and Hedgehog signaling [21–24]. Among these, the Hedgehog (Hh) signaling pathway is an evolutionarily conserved signaling cascade involved in embryonic development, tissue maintenance, and tumorigenesis [25,26]. Canonical Hh signaling is activated when Hedgehog ligands bind to Patched1 (PTCH1), thereby relieving Smoothened (SMO) inhibition and inducing downstream GLI family zinc finger 1 (GLI1)-dependent transcription [27–29]. Dysregulated Hh signaling has been associated with tumor growth, invasion, and malignant progression in multiple cancers [30,31], suggesting that it may also contribute to the aggressive behavior of EOCRC.

Several genes implicated in tumor progression may be relevant to EOCRC biology. Cell migration-inducing protein (CEMIP), also known as KIAA1199 or HYBID, has been linked to hyaluronic acid depolymerization, tumor progression, epithelial–mesenchymal transition (EMT), and metastasis [32]. CEMIP is upregulated during the transition from normal colorectal epithelium to adenomatous and neoplastic lesions and remains elevated in colorectal adenocarcinoma [33–35]. NKD2, a negative regulator of Wnt signaling, and FOXQ1, a Forkhead box transcription factor associated with EMT, invasion, and metastasis, have also been implicated in CRC [36–50]. However, whether these genes are enriched

in age-defined EOCRC cohorts remains unclear. We therefore integrated public transcriptomic datasets to identify EOCRC-associated candidates and used survival and in vitro loss-of-function analyses to prioritize and characterize CEMIP. Because pathway enrichment suggested Hedgehog signaling, we additionally examined the association of CEMIP with PTCH1, SMO, and GLI1. This pathway analysis was designed as hypothesis-generating rather than as evidence of a direct molecular interaction.

2. Materials and Methods

2.1. Public Dataset Analysis

Publicly available gene-expression datasets were retrieved from the NCBI Gene Expression Omnibus (GEO). GSE21815 was used as the discovery cohort, with cases stratified as EOCRC (<50 years) or LOCRC (≥50 years) according to the deposited age-at-diagnosis metadata. Ten previously identified CRC-associated genes were compared between the two groups. Integrated validation was performed using GSE20970, GSE36335, GSE37892, GSE45404, GSE56386, GSE73360, GSE89076, GSE101896, GSE106582, and GSE161158. For pathway analysis, GSE39575, comprising three control and three CEMIP-overexpressing SW480 samples, was analyzed; genes with nominal $p < 0.05$ were submitted for enrichment analysis. Detailed histological grade, complete TNM categories, molecular subtype, and treatment variables were not consistently available across the deposited metadata and were therefore not used for clinicopathological subgroup analyses.

2.2. Survival and Gene-Correlation Analyses

Overall survival associations for NKD2, FOXQ1, and CEMIP were assessed using Kaplan–Meier Plotter in the colorectal cancer cohort (Department of Bioinformatics, Semmelweis University, Budapest, Hungary; accessed in June 2026) [51]. Patients were separated into high- and low-expression groups using the platform-defined optimal cutoff, and hazard ratios (HRs), 95% confidence intervals (CIs), and log-rank p values were recorded. Gene-expression correlations between CEMIP and PTCH1, SMO, or GLI1 were evaluated using ENCORI/starBase in The Cancer Genome Atlas colon adenocarcinoma (TCGA-COAD) cohort ($n = 471$) [52]. The coefficients and p values generated by the platform were reported. Because complete individual-level clinicopathological and molecular covariates were not available through these interfaces for all analyzed cases, multivariable and age-specific survival subgroup analyses were not performed.

2.3. Enrichment Analysis and Visualization

Genes associated with CEMIP overexpression in GSE39575 were analyzed using Enrichr against the MSigDB Hallmark 2020 collection. Enrichment results were visualized using ImageGP. Nominal p values and multiple-testing-adjusted p values were retained, and the complete output is provided in Supplementary Table S1. Pathways were considered hypothesis-generating and were not interpreted as evidence of direct protein–protein interaction or pathway dependence.

2.4. Cell Lines, Culture Conditions, and Reagents

The selection of colorectal cancer cell lines was guided by the age of the patients from whom the original tumors were derived, based on the molecular characterization reported by Medico et al. [53]. Among the colorectal cancer cell lines with available patient-age information, HCT116 and HT29 were derived from patients younger than 50 years and were therefore used as EOCRC-derived models, whereas HCT15 and WiDr were derived from patients aged 50 years or older and were used as LOCRC-derived models. CRL-1459 was included as a non-malignant colon-derived reference cell line. HCT116 and HT29

were further selected for downstream functional analyses because they exhibited relatively high endogenous CEMIP expression. Cells were cultured in RPMI-1640 medium (Corning Incorporated, Corning, NY, USA) supplemented with 10% fetal bovine serum (FBS) (Corning Incorporated, Corning, NY, USA) and 1% penicillin-streptomycin (Thermo Fisher Scientific, Waltham, MA, USA; final concentrations: 50 U/mL penicillin and 50 µg/mL streptomycin). Cells were maintained at 37 °C in a humidified incubator containing 5% CO₂.

2.5. Cell Transfection

For loss-of-function experiments, a CEMIP-targeting siRNA and a general non-targeting negative-control siRNA were purchased from Invitrogen (Carlsbad, CA, USA). The control siRNA was not designed as a sequence-matched control specifically for the CEMIP-targeting siRNA; it was used to account for nonspecific effects of oligonucleotide delivery and transfection. The exact siRNA sequences were not available in the archived experimental records and therefore are not reported. HCT116 and HT29 cells were transfected with 50 or 100 nM siRNA using jetPRIME transfection reagent (Polyplus-transfection, New York, NY, USA) according to the manufacturer's instructions. Knockdown efficiency was assessed at the mRNA level by RT-qPCR 24 h after transfection.

2.6. Quantitative Reverse-Transcription PCR (RT-qPCR)

Total RNA was extracted using the RNeasy Kit and QIAshredder columns (Qiagen, Valencia, CA, USA), and 1 µg of RNA was reverse-transcribed using the iScript cDNA Synthesis Kit (Bio-Rad, Hercules, CA, USA). Quantitative PCR was performed with Qiagen SYBR Green reagent using the following program: 95 °C for 3 min; 45 cycles of 95 °C for 5 s, 55 °C for 20 s, and 72 °C for 5 s; followed by melt-curve analysis. Each sample was analyzed in triplicate. Transcript abundance was quantified using the standard-curve method and normalized to 18S rRNA. The primer sequences (5'–3') used in this study were as follows: CEMIP, forward GCTCTTGAGTTGCATGGACA and reverse ACCGCTTCAAATACTGGAC; NKD2, forward AACACGGAGCGCAGAAAC and reverse GGCC-GAATCTGGACGTGTA; FOXQ1, forward GCAAGTTCCCCTTTTCC and reverse AGTCGTTGAGCGAAAGGTT; PTCH1, forward TGCTGTGCCTGTGGTCATCCTGATT and reverse CAGAGCGAG-CATAGCCCTGTGGTTC; SHH, forward GATGTCTGCTGCTAGTCCTCG and reverse CACCTCTGAGTCATCAGCCTG; SMO, forward TGGCCG-CTTCAGATGACAGATGTTG and reverse CGTTAGCCGAATGTCAGCCGTGAAG; GLI1, forward CCCATAGGGTCTCGGGGTCTCAAAC and reverse GGAGGAC-CTGCGGCTGACTGTGTAA; and 18S, forward CATGGCCGTTCTTAGTTGGTGG and reverse CGCTGAGCCAGTCAGTGTAG.

2.7. Western Blot Analysis

Basal CEMIP expression was examined in the cell-line panel. Cells were lysed in buffer containing protease and phosphatase inhibitors, and lysates were centrifuged at 12,000 rpm for 10 min. Protein concentrations were determined using a Bio-Rad protein assay with bovine serum albumin as the standard. Equal amounts of protein (30 µg) were separated by SDS-PAGE, transferred to PVDF membranes, blocked with 5% bovine serum albumin, and incubated with primary antibodies overnight at 4 °C. After incubation with secondary antibodies (1:10,000), bands were detected by enhanced chemiluminescence using an Amersham Imager 600 (GE Healthcare Bio-Sciences AB, Uppsala, Sweden) and quantified with ImageJ (National Institutes of Health, Bethesda, MD, USA). Protein-level knockdown was not assessed in the archived experiments.

2.8. Cell Viability Assay

Relative viable cell biomass was measured by crystal violet staining. At 24 h after siRNA transfection, cells were seeded in 96-well plates and cultured for an additional 24 h. Cells were stained with 0.5% crystal violet for 10 min, washed, dried, and solubilized in 0.1 M sodium citrate. Absorbance was measured at 550 nm and expressed relative to the control siRNA group. This assay estimates adherent viable cell biomass and was not interpreted as a direct Ki67- or cell-cycle-based proliferation measurement.

2.9. Transwell Migration Assay

Cell migration was assessed using Transwell chambers with 8- μ m pore-size membranes. At 24 h after control or CEMIP siRNA transfection, 5×10^5 HCT116 or HT29 cells in serum-free medium were seeded in the upper chamber; the lower chamber contained medium supplemented with 10% fetal bovine serum as a chemoattractant. After 12 h, cells that had traversed the membrane and reached its lower surface were fixed with methanol, stained with Giemsa, imaged, and quantified. The blue-purple staining in the representative micrographs identifies migrated Giemsa-positive cells.

2.10. Wound-Healing Assay

The effect of CEMIP on cell migration was further assessed using a wound-healing assay kit (ibidi, Gräfelfing, Germany). HCT116 and HT29 cells were seeded at 4×10^5 cells/well. After 24 h, the inserts were removed to generate a wound gap, and images were captured at 0, 8, 12, and 24 h. The wound-healing rate was quantified using ImageJ.

2.11. Statistical Analysis

Cell-based experiments were performed in three independent biological replicates and are presented as the mean $\pm$ standard deviation. Two-group comparisons were performed using an unpaired two-tailed Student's *t*-test in GraphPad Prism 8.0. A *p* value < 0.05 was considered statistically significant. Survival analyses used the log-rank test and HRs generated by Kaplan–Meier Plotter. Correlation coefficients and *p* values were obtained from ENCORI. Enrichment results are reported with both nominal and adjusted *p* values; only pathways meeting the database-adjusted significance threshold were described as statistically enriched.

2.12. Generative Artificial Intelligence Disclosure

Generative artificial intelligence (GenAI), specifically ChatGPT (OpenAI, GPT-5.6 Thinking), was used only for language refinement and manuscript editing. No GenAI tools were used for data analysis, data interpretation, figure generation, or scientific decision-making.

3. Results

3.1. Integrated Transcriptomic Analyses Identify CEMIP, NKD2, and FOXQ1 as EOCRC-Associated Candidates

Based on our previous study, ten genes (DPEP1, KRT80, FABP6, NKD2, FOXQ1, CEMIP, ETV4, TESC, FUT1, and GAS2) were selected for age-stratified evaluation [54]. In GSE21815, CEMIP, NKD2, and FOXQ1 were significantly higher in the EOCRC group than in the LOCRC group (Figure 1A). The available clinical characteristics of the discovery cohort are summarized in Table 1. Integrated analysis of ten independent GEO datasets showed the same direction of differential expression for these three genes (Figure 1B). Detailed histological grade, complete TNM categories, molecular subtype, and treatment information were not consistently available in the deposited metadata; therefore,

robust clinicopathological correlation analyses could not be performed. These results identify CEMIP, NKD2, and FOXQ1 as candidate genes associated with age-defined EOCRC cohorts.

Table 1. Early-onset and late-onset colon cancer details in CRC Clinic. (GSE21815, Japanese CRC patient profile).

Characteristic	EOCRC		LOCRC	
	Number	%	Number	%
Number of patients	14	100%	123	100%
Gender				
Male	10	71%	68	55%
Female	4	29%	55	45%
Age at diagnosis (years)				
31–40	4	29%	0	0
41–50	10	71%	0	0
51–60	0	0	35	28%
61–70	0	0	42	33%
71–80	0	0	38	22%
>80	0	0	8	7%
TNM stage				
1	0	0	12	10%
2	2	15%	25	20%
3A	0	0	11	9%
3B	0	0	5	4%
4	3	21%	8	7%
ND	9	64%	62	50%

3.2. Overall Survival Analysis Prioritizes CEMIP for Further Study

To prioritize the three EOCRC-associated candidates, we evaluated their associations with overall survival in an independent colorectal cancer survival resource. High CEMIP expression was significantly associated with poorer survival (HR = 1.40, 95% CI = 1.09–1.79; log-rank $p = 0.008$), whereas NKD2 showed a borderline association (HR = 1.25, 95% CI = 0.98–1.60; $p = 0.067$) and FOXQ1 was not significant (HR = 1.19, 95% CI = 0.92–1.54; $p = 0.18$) (Figure 2). The numbers displayed beneath each Kaplan–Meier curve constitute the number-at-risk table and indicate how many patients in the low- and high-expression groups remained under observation at 0, 50, 100, 150, and 200 months. These overall CRC survival data were used for candidate prioritization and do not establish an EOCRC-specific prognostic effect. Clinicopathological and molecular subgroup analyses were not performed because complete individual-level covariates were not available through the survival interface. On the basis of its reproducible transcriptomic enrichment and significant overall survival association, CEMIP was selected for subsequent functional analyses.

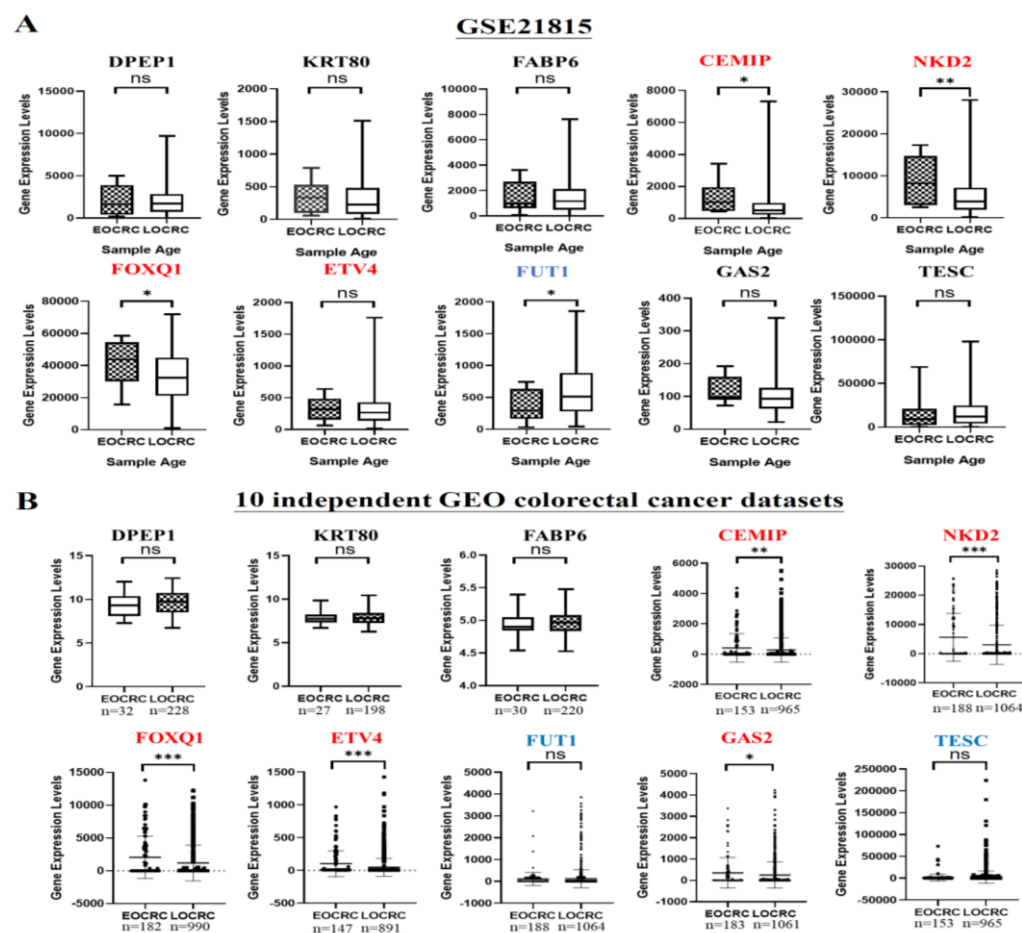


Figure 1. Integrated expression analysis of ten CRC-associated genes in EOCRC and LOCRC. (A) Age-stratified expression in the GSE21815 discovery cohort. CEMP, NKD2, and FOXQ1 were significantly higher in EOCRC. (B) Integrated analysis of ten independent GEO colorectal cancer datasets showing the same direction of differential expression for CEMP, NKD2, and FOXQ1. Red labels indicate genes higher in EOCRC; blue labels indicate genes higher in LOCRC; ns: non-significant. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

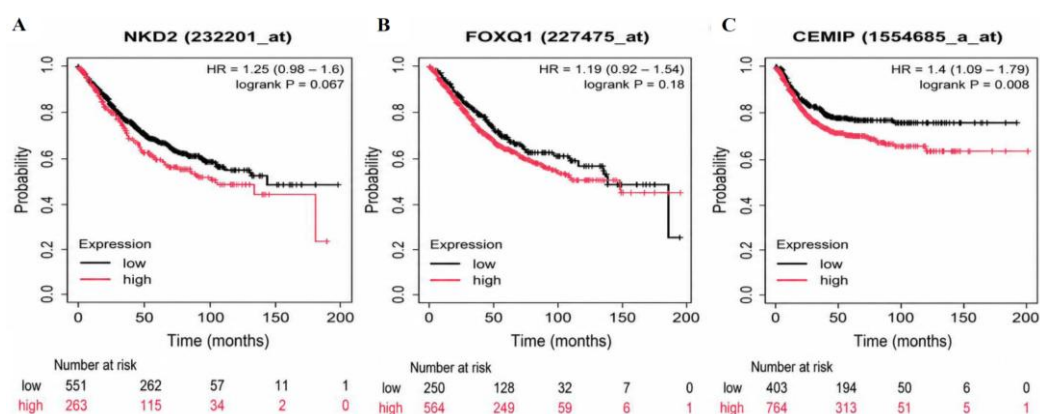


Figure 2. Overall survival analysis of NKD2, FOXQ1, and CEMP in colorectal cancer. Kaplan–Meier curves compare high- and low-expression groups for NKD2 (A), FOXQ1 (B), and CEMP (C). The number-at-risk tables beneath the curves report the numbers of patients remaining under observation in the low-expression (black) and high-expression (red) groups at 0, 50, 100, 150, and 200 months. Only CEMP was significantly associated with poorer overall survival (HR = 1.40, 95% CI = 1.09–1.79; log-rank $p = 0.008$). These analyses were performed in an overall colorectal cancer cohort and were used for candidate prioritization rather than for EOCRC-specific prognostic inference.

3.3. CEMIP Is Highly Expressed in Selected CRC Cell Lines and Supports Viable Cell Biomass

We first examined the basal expression of CEMIP in the non-malignant colon-derived CCD-18Co cell line (ATCC CRL-1459) and four colorectal cancer cell lines, including HCT116, HT29, HCT15, and WiDr. RT-qPCR analysis showed marked variation in basal CEMIP mRNA expression among the tested cell models. HCT116 cells exhibited the highest CEMIP mRNA level, followed by HT29 cells, whereas substantially lower expression was detected in CCD-18Co, HCT15, and WiDr cells (Figure 3A). Consistent with the transcript-level findings, Western blot analysis demonstrated strong CEMIP expression in HCT116 and HT29 cells but only weak or minimal expression in the remaining cell lines (Figure 3B). Densitometric analysis normalized to β -actin confirmed the relatively high abundance of CEMIP in HCT116 and HT29 cells (Figure 3C). Based on these concordant mRNA and protein expression profiles, HCT116 and HT29 cells were selected for subsequent loss-of-function experiments. The efficiency of siRNA-mediated CEMIP silencing was subsequently evaluated by RT-qPCR. Relative CEMIP mRNA expression, normalized to 18S rRNA, was significantly reduced following CEMIP siRNA transfection in both HCT116 and HT29 cells compared with the corresponding control siRNA groups (Figure 3D,F). The reduction was pronounced in both models, with particularly strong suppression observed in HT29 cells. We next assessed the effect of CEMIP knockdown on viable adherent cell biomass using crystal violet staining. CEMIP silencing caused a significant and concentration-dependent reduction in relative viable cell biomass in both cell lines (Figure 3E,G). In HCT116 cells, treatment with 50 and 100 nM CEMIP siRNA reduced the crystal violet signal relative to control siRNA, with the greatest reduction observed at 100 nM. A similar but more pronounced dose-dependent response was observed in HT29 cells. These findings indicate that endogenous CEMIP contributes to the maintenance of viable adherent CRC cell biomass, although the assay does not distinguish among effects on proliferation, survival, and cell attachment.

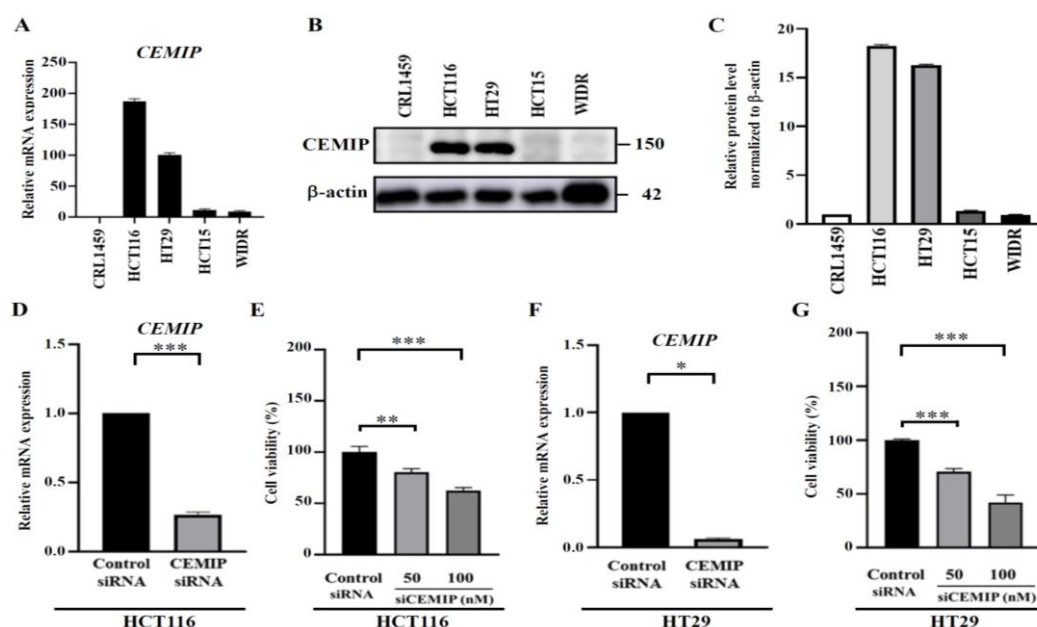


Figure 3. Basal CEMIP expression and effects of CEMIP knockdown on viable cell biomass. (A) Basal CEMIP mRNA expression in CCD-18Co, HCT116, HT29, HCT15, and WiDr cells. (B) Representative Western blot of basal CEMIP expression. (C) Densitometric quantification normalized to β -actin. (D,F) RT-qPCR confirmation of CEMIP mRNA knockdown in HCT116 and HT29 cells. (E,G) Crystal violet assays showing reduced relative viable cell biomass after treatment with 50 or 100 nM CEMIP

siRNA. Panels D–G were horizontally aligned, and the CEMIP-targeting reagent is labeled consistently as CEMIP siRNA. Data are mean $\pm$ SD from three independent experiments. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

3.4. CEMIP Knockdown Reduces the Migratory Capacity of CRC Cells

We next investigated whether CEMIP contributes to CRC cell migration using complementary wound-healing and Transwell assays. In the wound-healing assay, the initial wound area was defined as 100% at 0 h for each experimental group. Control HCT116 cells showed progressive wound closure, with the remaining wound area decreasing to approximately 89% at 8 h and 54% at 12 h, followed by nearly complete closure at 24 h. In contrast, CEMIP-silenced HCT116 cells showed delayed wound closure, retaining approximately 95%, 83%, and 63% of the initial wound area at 8, 12, and 24 h, respectively (Figure 4A). A similar pattern was observed in HT29 cells. The remaining wound area in control cells decreased to approximately 80% at 8 h and 58% at 12 h, with nearly complete closure by 24 h. However, HT29 cells transfected with CEMIP siRNA retained approximately 90%, 81%, and 57% of the initial wound area at the corresponding time points (Figure 4A). These findings indicate that CEMIP knockdown consistently delays wound closure in both CRC cell models. The effect of CEMIP silencing on migration was further examined using the Transwell assay. Representative micrographs showed fewer migrated, Giemsa-stained cells after CEMIP knockdown in both HCT116 and HT29 cells compared with their respective control groups (Figure 4B). Quantification confirmed a significant reduction in migrated cell numbers following CEMIP silencing (Figure 4C). The reduction was particularly marked in HCT116 cells, in which the number of migrated cells was decreased by approximately one-half, whereas HT29 cells showed a smaller but statistically significant reduction. Taken together, the results from both wound-healing and Transwell assays consistently demonstrate that CEMIP knockdown impairs the migratory capacity of CRC cells. These data support a positive contribution of endogenous CEMIP to CRC cell motility, while the potential contribution of reduced viable cell biomass to the migration phenotype should also be considered.

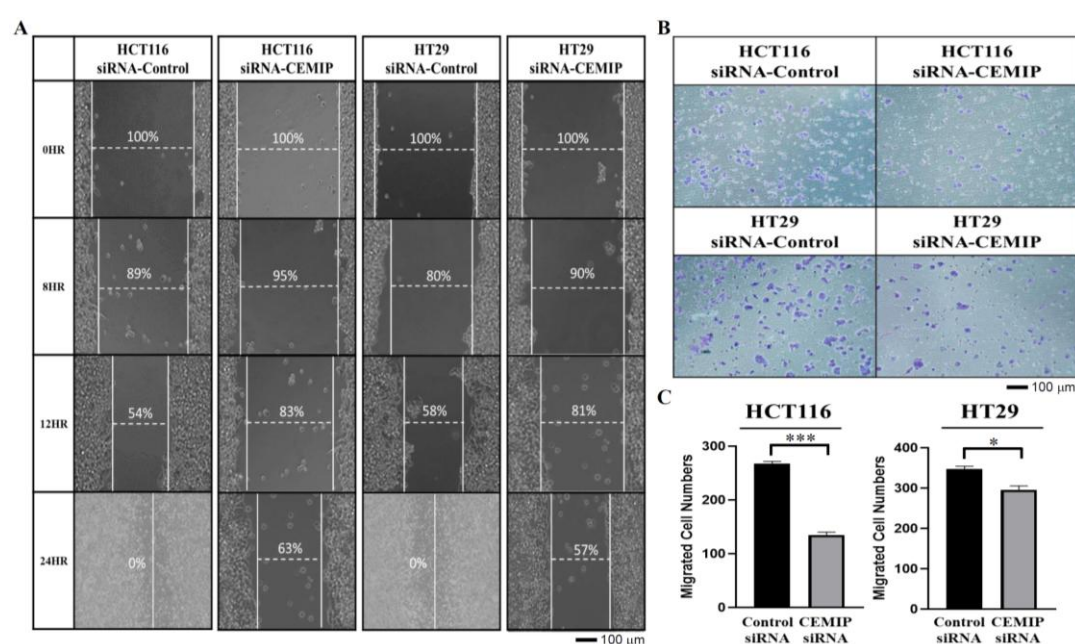


Figure 4. CEMIP knockdown reduces CRC cell migration. (A) Representative wound-healing images and wound-closure measurements in HCT116 and HT29 cells following transfection with control siRNA or CEMIP siRNA. (B) Representative Transwell migration images showing Giemsa-stained cells that migrated through the 8- μ m-pore membrane and adhered to its lower surface. (C)

Quantification of migrated cells. One representative scale bar is shown for each microscopy assay panel because all images within the corresponding panel were acquired at the same magnification. Scale bars = 100 μ m. Data are presented as the mean $\pm$ SD from three independent experiments. The HCT116 comparison was significant at *** $p < 0.001$, whereas the HT29 comparison was significant at * $p < 0.05$.

3.5. CEMIP Expression Is Associated with PTCH1-Related Hedgehog Signaling

To identify signaling pathways potentially associated with CEMIP, genes differentially expressed in CEMIP-overexpressing SW480 cells from the GSE39575 dataset were subjected to enrichment analysis using the MSigDB Hallmark gene sets. Hedgehog signaling was among the significantly enriched pathways, together with apical junction, estrogen-response, hypoxia, and mTORC1-related signatures. In contrast, Wnt/ β -catenin and TGF- β signaling showed nominal enrichment but did not remain significant after adjustment for multiple comparisons (Figure 5A; Supplementary Table S1). We next examined the expression of selected Hedgehog-related genes after CEMIP knockdown in HCT116 and HT29 cells. The analyzed genes included Patched 1 (PTCH1), Smoothened (SMO), and GLI family zinc finger 1 (GLI1). Relative mRNA levels were normalized to 18S rRNA and expressed relative to the corresponding control siRNA group. In both cell lines, CEMIP silencing significantly reduced PTCH1 and GLI1 mRNA expression, whereas SMO mRNA expression was increased (Figure 5B,C). To further assess these relationships in clinical colorectal cancer samples, we analyzed TCGA colon adenocarcinoma (TCGA-COAD) RNA-sequencing data. Gene expression values were reported as log2-transformed FPKM values [$\log_2(\text{FPKM} + 0.01)$]. CEMIP expression showed a moderate positive correlation with PTCH1 ($r = 0.345$, $p = 1.31 \times 10^{-14}$), a weak positive correlation with SMO ($r = 0.108$, $p = 1.88 \times 10^{-2}$), and no significant correlation with GLI1 ($r = 0.032$, $p = 0.485$) (Figure 5D–F). Taken together, these findings support an exploratory association between CEMIP expression and PTCH1-related Hedgehog signaling. However, the current data do not establish a direct molecular interaction, pathway dependency, or a linear CEMIP–PTCH1–SMO–GLI1 signaling cascade.

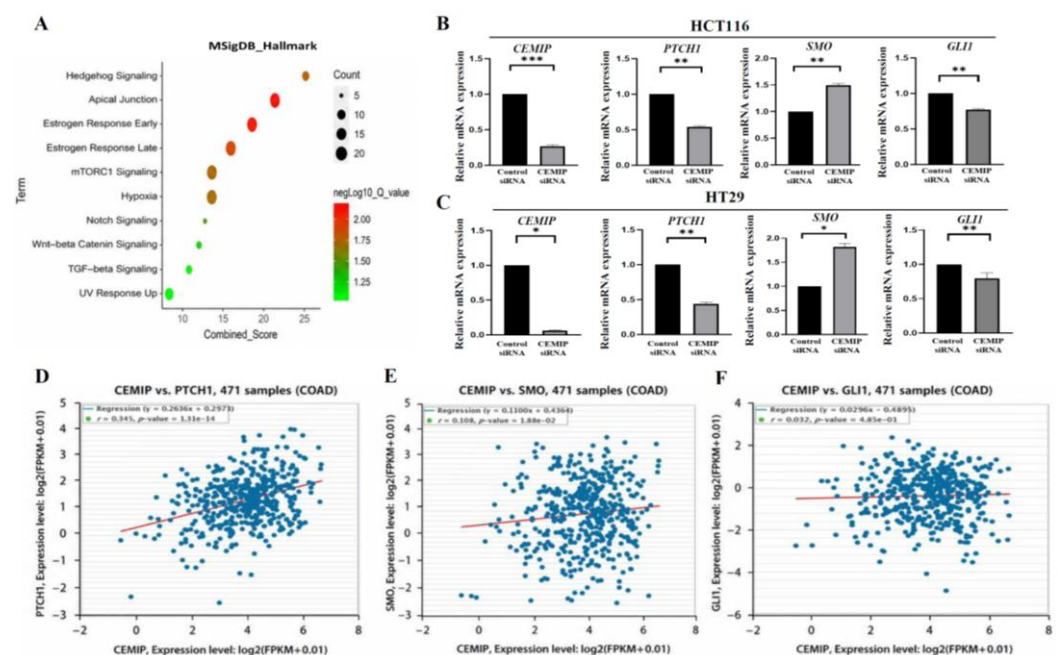


Figure 5. Exploratory association of CEMIP with PTCH1-related Hedgehog signaling. (A) MSigDB Hallmark 2020 enrichment analysis of genes associated with CEMIP overexpression in GSE39575. (B,C) RT-qPCR analysis of CEMIP, PTCH1, SMO, and GLI1 after CEMIP knockdown in HCT116

and HT29 cells; transcript levels were normalized to 18S rRNA and expressed relative to the corresponding control-siRNA group. (D–F) Redrawn correlation plots showing ENCORI associations of CEMIP with PTCH1 ($r = 0.345$, $p = 1.31 \times 10^{-14}$), SMO ($r = 0.108$, $p = 1.88 \times 10^{-2}$), and GLI1 ($r = 0.032$, $p = 0.485$) in TCGA-COAD samples. Axes report $\log_2(\text{FPKM} + 0.01)$ expression values. The data support an association but not direct molecular interaction or causality. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

4. Discussion

In this study, integrated age-stratified transcriptomic analyses identified CEMIP, NKD2, and FOXQ1 as candidate genes enriched in EOCRC cohorts. Overall CRC survival analysis prioritized CEMIP, and loss-of-function experiments showed that CEMIP supports viable cell biomass and migration in HCT116 and HT29 cells. Pathway enrichment and gene-expression analyses additionally revealed an exploratory association between CEMIP and PTCH1-related Hedgehog signaling. Importantly, the revised interpretation distinguishes functional evidence for CEMIP in the measured cellular phenotypes from correlative evidence regarding the Hedgehog pathway.

EOCRC is increasingly viewed as a heterogeneous clinical entity rather than simply CRC occurring at a younger age. Systematic and case-controlled studies have reported differences in presentation, histology, and molecular profiles between EOCRC and LOCRC [55–57]. Our discovery cohort was small and incompletely annotated, but the direction of CEMIP, NKD2, and FOXQ1 expressions were reproduced in an integrated set of independent GEO cohorts. The survival analysis was conducted in an overall CRC population and therefore served only to prioritize CEMIP; it does not prove that CEMIP has an age-specific prognostic effect.

The functional phenotype observed after CEMIP knockdown is consistent with established CRC biology. CEMIP/KIAA1199 is induced during colorectal neoplastic progression and is associated with poor outcome [58]. Mechanistically, CEMIP has been linked to glutamine metabolic reprogramming and β -catenin regulation [59], MHC-I internalization and immune evasion [60], Wnt/ β -catenin/Snail-dependent EMT [61], and GRAF1–MIB1–CDC42/MAPK-mediated EMT and metastasis [62]. These pathways provide a stronger established mechanistic context for the reduced cell viability and migration observed here than the currently exploratory Hedgehog association.

Recent studies further broaden the relevance of CEMIP to stemness and inflammatory tumor ecology. The ELK4–METTL3–CEMIP axis was reported to enhance colorectal carcinoma cell growth, migration, invasion, spheroid formation, and stemness [63]. In addition, an inflammation-associated tumor-microenvironment analysis identified CEMIP-positive monocytes as an immunosuppressive subgroup linked to unfavorable anti-PD-1 response [64], complementing evidence that tumor-derived CEMIP can reduce MHC-I surface presentation [60]. These observations suggest that future EOCRC studies should evaluate CEMIP in tumor epithelial cells and immune/stromal compartments rather than assuming a tumor-cell-autonomous mechanism.

Our pathway analysis requires cautious interpretation. STRING does not support a direct CEMIP–PTCH1 protein interaction, and no complementary PTCH1 knockdown, rescue, ligand stimulation, or pathway-inhibition experiments were performed. Moreover, the correlation with PTCH1 was moderate, the correlation with SMO was weak, and GLI1 was not correlated in TCGA-COAD. Accordingly, the data support an association with PTCH1-related Hedgehog signaling rather than a direct mechanistic axis. The decrease in PTCH1 and GLI1 together with an increase in SMO after CEMIP silencing may reflect compensatory feedback, altered receptor trafficking, or non-canonical Hedgehog

regulation. Because transcript abundance does not directly establish pathway activity, future studies should measure GLI-dependent transcription, protein localization, and pharmacologic or genetic rescue.

From a translational perspective, CEMIP may be useful as part of a molecular stratification panel for CRC, particularly if its age-associated expression is validated in larger clinically annotated EOCRC cohorts. Direct drug development against CEMIP remains challenging because it is a multifunctional extracellular and intracellular protein; however, its established links to CDC42/MAPK signaling, metabolic adaptation, stemness, and immune evasion suggest rational combination strategies. Potential approaches include testing CEMIP-directed agents or pathway inhibitors alongside immune-checkpoint blockade, metabolic therapies, or agents that target EMT/stemness programs. By contrast, combining CEMIP targeting with Hedgehog inhibitors would be premature until direct pathway dependence is demonstrated.

This study has several limitations. First, the EOCRC discovery group was small and imbalanced relative to the LOCRC group, and the available public metadata did not support comprehensive age-specific survival or clinicopathological subgroup analyses. Second, HCT116 and HT29 were selected as EOCRC-derived models according to the reported age of the patients from whom the tumors originated and because of their relatively high endogenous CEMIP expression; however, two established cell lines cannot reproduce the biological and clinical heterogeneity of EOCRC, and age at tumor derivation alone does not establish an EOCRC-specific cellular phenotype. Third, the functional experiments were limited to siRNA-mediated loss of function without sequence disclosure, protein-level knockdown confirmation, rescue experiments, patient-derived models, or in vivo validation. Finally, the PTCH1/Hedgehog findings are based on enrichment, transcript-level perturbation, and correlation analyses and should therefore remain hypothesis-generating.

Future investigations should validate CEMIP expression in independent EOCRC tissue cohorts with detailed TNM stage, grade, molecular subtype, treatment, and outcome data; incorporate patient-derived organoids and in vivo models; and test whether CEMIP perturbation alters EMT, inflammasome/inflammatory signaling, stemness, immune evasion, and Hedgehog pathway activity. Such studies will determine whether CEMIP is primarily a biomarker, a context-dependent functional contributor, or a tractable therapeutic target in EOCRC.

5. Conclusions

Integrated transcriptomic analyses identified CEMIP, NKD2, and FOXQ1 as candidate genes associated with age-defined EOCRC cohorts, and overall CRC survival analysis prioritized CEMIP for functional evaluation. CEMIP knockdown reduced relative viable cell biomass and migration in CRC cells. CEMIP also showed an exploratory association with PTCH1-related Hedgehog signaling; however, the available data do not establish a direct interaction or pathway-mediated causality. CEMIP should therefore be considered a candidate EOCRC biomarker and functional contributor whose mechanistic and translational relevance requires validation in clinically annotated cohorts and direct pathway-modulation models.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/genes17080922/s1>, Table S1: The MSigDB_Hallmark_2020_table analysis comes from the GSE39575 data set, and the correlation analysis with the CEMIP expression.

Author Contributions: Conceptualization, Y.-F.W. and K.-H.L.; methodology, C.-Y.L., H.-J.P., and H.D.N.; validation, C.-Y.L., H.-J.P., Y.-C.L., Y.-C.N., K.X.Y., and M.T.N.; formal analysis, C.-Y.L., H.-J.P., Y.-C.L., and Y.-C.N.; investigation, C.-Y.L., H.-J.P., H.D.N., Y.-C.N., K.X.Y., and M.T.N.; resources, Y.-F.W. and K.-H.L.; data curation, C.-Y.L. and H.-J.P.; writing—original draft preparation, C.-Y.L. and H.-J.P.; writing—review and editing, Y.-F.W. and K.-H.L.; visualization, C.-Y.L. and H.-J.P.; supervision, Y.-F.W. and K.-H.L.; project administration, Y.-F.W. and K.-H.L.; funding acquisition, Y.-F.W. and K.-H.L. C.-Y.L. and H.-J.P. contributed equally to this work. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Tri-Service General Hospital (grant number: TSGH-E-114295, TSGH-E-115297, awarded to Y.-F.W.); the Taipei Medical University–Ministry of Health and Welfare Shuang Ho Hospital Featured Research Program (grant number: 112FRP-27, awarded to C.-Y.L.); and the National Science and Technology Council (grant numbers: NSTC114-2320-B-038-006, NSTC 114-2622-B-038-007-HONG JING CO. and NSTC115-2320-B-038-066-MY3, awarded to K.-H.L.).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Publicly available gene-expression datasets were obtained from the NCBI Gene Expression Omnibus under accession numbers GSE21815, GSE39575, GSE20970, GSE36335, GSE37892, GSE45404, GSE56386, GSE73360, GSE89076, GSE101896, GSE106582, and GSE161158. Survival data were analyzed using Kaplan–Meier Plotter, and gene-correlation data were analyzed using ENCORI/starBase. The processed data underlying the figures are available from the corresponding authors upon reasonable request.

Acknowledgments: During preparation of the revised manuscript, the authors used ChatGPT (OpenAI, GPT-5.6 Thinking) for language refinement, grammar correction, sentence restructuring, and improvement of manuscript clarity. No generative artificial intelligence tool was used to generate experimental data, perform statistical analyses, create figures, or make scientific decisions. The authors reviewed and edited all outputs and take full responsibility for the publication content.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

EOCRC	Early-onset colorectal cancer
CEMIP	Cell migration-inducing protein
Hh	Hedgehog
PTCH1	Patched 1

References

1. Xi, Y.; Xu, P. Global colorectal cancer burden in 2020 and projections to 2040. *Transl. Oncol.* **2021**, *14*, 101174. <https://doi.org/10.1016/j.tranon.2021.101174>.
2. Fleming, M.; Ravula, S.; Tatishchev, S.F.; Wang, H.L. Colorectal carcinoma: Pathologic aspects. *J. Gastrointest. Oncol.* **2012**, *3*, 153–173. <https://doi.org/10.3978/j.issn.2078-6891.2012.030>.
3. Jaspersion, K.W.; Tuohy, T.M.; Neklason, D.W.; Burt, R.W. Hereditary and familial colon cancer. *Gastroenterology* **2010**, *138*, 2044–2058. <https://doi.org/10.1053/j.gastro.2010.01.054>.
4. Jiao, S.; Peters, U.; Berndt, S.; Brenner, H.; Butterbach, K.; Caan, B.J.; Carlson, C.S.; Chan, A.T.; Chang-Claude, J.; Chanock, S.; et al. Estimating the heritability of colorectal cancer. *Hum. Mol. Genet.* **2014**, *23*, 3898–3905. <https://doi.org/10.1093/hmg/ddu087>.
5. Conteduca, V.; Sansonno, D.; Russi, S.; Dammacco, F. Precancerous colorectal lesions (Review). *Int. J. Oncol.* **2013**, *43*, 973–984. <https://doi.org/10.3892/ijo.2013.2041>.

6. Kedrin, D.; Gala, M.K. Genetics of the serrated pathway to colorectal cancer. *Clin. Transl. Gastroenterol.* **2015**, *6*, e84. <https://doi.org/10.1038/ctg.2015.12>.
7. Jess, T.; Rungoe, C.; Peyrin-Biroulet, L. Risk of colorectal cancer in patients with ulcerative colitis: A meta-analysis of population-based cohort studies. *Clin. Gastroenterol. Hepatol.* **2012**, *10*, 639–645. <https://doi.org/10.1016/j.cgh.2012.01.010>.
8. Triantafyllidis, J.K.; Nasioulas, G.; Kosmidis, P.A. Colorectal cancer and inflammatory bowel disease: Epidemiology, risk factors, mechanisms of carcinogenesis and prevention strategies. *Anticancer. Res.* **2009**, *29*, 2727–2737.
9. Venugopal, A.; Carethers, J.M. Epidemiology and biology of early onset colorectal cancer. *EXCLI J.* **2022**, *21*, 162–182. <https://doi.org/10.17179/excli2021-4456>.
10. 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. <https://doi.org/10.1002/1878-0261.12417>.
11. Patel, S.G.; Ahnen, D.J. Colorectal Cancer in the Young. *Curr. Gastroenterol. Rep.* **2018**, *20*, 15. <https://doi.org/10.1007/s11894-018-0618-9>.
12. Connell, L.C.; Mota, J.M.; Braghiroli, M.I.; Hoff, P.M. The Rising Incidence of Younger Patients With Colorectal Cancer: Questions About Screening, Biology, and Treatment. *Curr. Treat. Options Oncol.* **2017**, *18*, 23. <https://doi.org/10.1007/s11864-017-0463-3>.
13. Hubbard, J.M.; Grothey, A. Adolescent and young adult colorectal cancer. *J. Natl. Compr. Cancer Netw.* **2013**, *11*, 1219–1225. <https://doi.org/10.6004/jnccn.2013.0144>.
14. Patel, S.G.; Boland, C.R. Colorectal Cancer in Persons Under Age 50: Seeking Causes and Solutions. *Gastrointest. Endosc. Clin. N. Am.* **2020**, *30*, 441–455. <https://doi.org/10.1016/j.giec.2020.03.001>.
15. 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. <https://doi.org/10.1007/s10620-019-05644-0>.
16. Kibriya, M.G.; Raza, M.; Quinn, A.; Kamal, M.; Ahsan, H.; Jasmine, F. A Transcriptome and Methylome Study Comparing Tissues of Early and Late Onset Colorectal Carcinoma. *Int. J. Mol. Sci.* **2022**, *23*, 14261. <https://doi.org/10.3390/ijms232214261>.
17. Vuik, F.E.; Nieuwenburg, S.A.; Bardou, M.; Lansdorp-Vogelaar, I.; Dinis-Ribeiro, M.; Bento, M.J.; Zadnik, V.; Pellisé, 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. <https://doi.org/10.1136/gutjnl-2018-317592>.
18. Sinicrope, F.A. Increasing Incidence of Early-Onset Colorectal Cancer. *N. Engl. J. Med.* **2022**, *386*, 1547–1558. <https://doi.org/10.1056/NEJMr2200869>.
19. Saraiva, M.R.; Rosa, I.; Claro, I. Early-onset colorectal cancer: A review of current knowledge. *World J. Gastroenterol.* **2023**, *29*, 1289–1303. <https://doi.org/10.3748/wjg.v29.i8.1289>.
20. Brabletz, T.; Kalluri, R.; Nieto, M.A.; Weinberg, R.A. EMT in cancer. *Nat. Rev. Cancer* **2018**, *18*, 128–134. <https://doi.org/10.1038/nrc.2017.118>.
21. Kumar, A.; Golani, A.; Kumar, L.D. EMT in breast cancer metastasis: An interplay of microRNAs, signaling pathways and circulating tumor cells. *FBL* **2020**, *25*, 979–1010. <https://doi.org/10.2741/4844>.
22. Kang, Y.; He, W.; Tulley, S.; Gupta, G.P.; Serganova, I.; Chen, C.R.; Manova-Todorova, K.; Blasberg, R.; Gerald, W.L.; Massagué, J. Breast cancer bone metastasis mediated by the Smad tumor suppressor pathway. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 13909–13914. <https://doi.org/10.1073/pnas.0506517102>.
23. Micalizzi, D.S.; Farabaugh, S.M.; Ford, H.L. Epithelial-mesenchymal transition in cancer: Parallels between normal development and tumor progression. *J. Mammary Gland Biol. Neoplasia* **2010**, *15*, 117–134. <https://doi.org/10.1007/s10911-010-9178-9>.
24. Katoh, Y.; Katoh, M. Hedgehog signaling, epithelial-to-mesenchymal transition and miRNA (review). *Int. J. Mol. Med.* **2008**, *22*, 271–275.
25. Varjosalo, M.; Taipale, J. Hedgehog: Functions and mechanisms. *Genes Dev.* **2008**, *22*, 2454–2472. <https://doi.org/10.1101/gad.1693608>.
26. Le, H.; Kleinerman, R.; Lerman, O.Z.; Brown, D.; Galiano, R.; Gurtner, G.C.; Warren, S.M.; Levine, J.P.; Saadeh, P.B. Hedgehog signaling is essential for normal wound healing. *Wound Repair Regen.* **2008**, *16*, 768–773. <https://doi.org/10.1111/j.1524-475X.2008.00430.x>.
27. Wu, C.; Zhu, X.; Liu, W.; Ruan, T.; Tao, K. Hedgehog signaling pathway in colorectal cancer: Function, mechanism, and therapy. *OncoTargets Ther.* **2017**, *10*, 3249–3259. <https://doi.org/10.2147/OTT.S139639>.
28. Mastronardi, F.G.; Dimitroulakos, J.; Kamel-Reid, S.; Manoukian, A.S. Co-localization of patched and activated sonic hedgehog to lysosomes in neurons. *Neuroreport* **2000**, *11*, 581–585. <https://doi.org/10.1097/00001756-200002280-00030>.

29. Rubin, L.L.; de Sauvage, F.J. Targeting the Hedgehog pathway in cancer. *Nat. Rev. Drug Discov.* **2006**, *5*, 1026–1033. <https://doi.org/10.1038/nrd2086>.
30. Huang, S.; He, J.; Zhang, X.; Bian, Y.; Yang, L.; Xie, G.; Zhang, K.; Tang, W.; Stelter, A.A.; Wang, Q.; et al. Activation of the hedgehog pathway in human hepatocellular carcinomas. *Carcinogenesis* **2006**, *27*, 1334–1340. <https://doi.org/10.1093/carcin/bgi378>.
31. Lauressergues, E.; Heusler, P.; Lestienne, F.; Troulier, D.; Rauly-Lestienne, I.; Tourette, A.; Ailhaud, M.C.; Cathala, C.; Tardif, S.; Denais-Laliève, D.; et al. Pharmacological evaluation of a series of smoothened antagonists in signaling pathways and after topical application in a depilated mouse model. *Pharmacol. Res. Perspect.* **2016**, *4*, e00214. <https://doi.org/10.1002/prp2.214>.
32. Spataro, S.; Guerra, C.; Cavalli, A.; Sgrignani, J.; Sleeman, J.; Poulain, L.; Boland, A.; Scapozza, L.; Moll, S.; Prunotto, M. CEMIP (HYBID, KIAA1199): Structure, function and expression in health and disease. *FEBS J.* **2022**, *290*, 3946–3962. <https://doi.org/10.1111/febs.16600>.
33. Sabates-Bellver, J.; Van der Flier, L.G.; de Palo, M.; Cattaneo, E.; Maake, C.; Rehrauer, H.; Laczko, E.; Kurowski, M.A.; Bujnicki, J.M.; Menigatti, M.; et al. Transcriptome profile of human colorectal adenomas. *Mol. Cancer Res.* **2007**, *5*, 1263–1275. <https://doi.org/10.1158/1541-7786.Mcr-07-0267>.
34. Birkenkamp-Demtroder, K.; Maghnooui, A.; Mansilla, F.; Thorsen, K.; Andersen, C.L.; Øster, B.; Hahn, S.; Ørntoft, T.F. Repression of KIAA1199 attenuates Wnt-signalling and decreases the proliferation of colon cancer cells. *Br. J. Cancer* **2011**, *105*, 552–561. <https://doi.org/10.1038/bjc.2011.268>.
35. LaPointe, L.C.; Pedersen, S.K.; Dunne, R.; Brown, G.S.; Pimlott, L.; Gaur, S.; McEvoy, A.; Thomas, M.; Wattchow, D.; Molloy, P.L.; et al. Discovery and validation of molecular biomarkers for colorectal adenomas and cancer with application to blood testing. *PLoS ONE* **2012**, *7*, e29059. <https://doi.org/10.1371/journal.pone.0029059>.
36. Katoh, M. Molecular cloning, gene structure, and expression analyses of NKD1 and NKD2. *Int. J. Oncol.* **2001**, *19*, 963–969.
37. Zeng, W.; Wharton, K.A., Jr.; Mack, J.A.; Wang, K.; Gadbaw, M.; Suyama, K.; Klein, P.S.; Scott, M.P. naked cuticle encodes an inducible antagonist of Wnt signalling. *Nature* **2000**, *403*, 789–795. <https://doi.org/10.1038/35001615>.
38. Rousset, R.; Mack, J.A.; Wharton, K.A., Jr.; Axelrod, J.D.; Cadigan, K.M.; Fish, M.P.; Nusse, R.; Scott, M.P. Naked cuticle targets dishevelled to antagonize Wnt signal transduction. *Genes Dev.* **2001**, *15*, 658–671. <https://doi.org/10.1101/gad.869201>.
39. Yan, D.; Wallingford, J.B.; Sun, T.Q.; Nelson, A.M.; Sakanaka, C.; Reinhard, C.; Harland, R.M.; Fantl, W.J.; Williams, L.T. Cell autonomous regulation of multiple Dishevelled-dependent pathways by mammalian Nkd. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 3802–3807. <https://doi.org/10.1073/pnas.071041898>.
40. Li, C.; Franklin, J.L.; Graves-Deal, R.; Jerome, W.G.; Cao, Z.; Coffey, R.J. Myristoylated Naked2 escorts transforming growth factor alpha to the basolateral plasma membrane of polarized epithelial cells. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 5571–5576. <https://doi.org/10.1073/pnas.0401294101>.
41. Götze, S.; Wolter, M.; Reifenberger, G.; Müller, O.; Sievers, S. Frequent promoter hypermethylation of Wnt pathway inhibitor genes in malignant astrocytic gliomas. *Int. J. Cancer* **2010**, *126*, 2584–2593. <https://doi.org/10.1002/ijc.24981>.
42. Dong, Y.; Cao, B.; Zhang, M.; Han, W.; Herman, J.G.; Fuks, F.; Zhao, Y.; Guo, M. Epigenetic silencing of NKD2, a major component of Wnt signaling, promotes breast cancer growth. *Oncotarget* **2015**, *6*, 22126–22138. <https://doi.org/10.18632/oncotarget.4244>.
43. Cao, B.; Yang, W.; Jin, Y.; Zhang, M.; He, T.; Zhan, Q.; Herman, J.G.; Zhong, G.; Guo, M. Silencing NKD2 by Promoter Region Hypermethylation Promotes Esophageal Cancer Progression by Activating Wnt Signaling. *J. Thorac. Oncol.* **2016**, *11*, 1912–1926. <https://doi.org/10.1016/j.jtho.2016.06.015>.
44. Tabatabai, R.; Linhares, Y.; Bolos, D.; Mita, M.; Mita, A. Targeting the Wnt Pathway in Cancer: A Review of Novel Therapeutics. *Target. Oncol.* **2017**, *12*, 623–641. <https://doi.org/10.1007/s11523-017-0507-4>.
45. Hu, T.; Krezel, A.M.; Li, C.; Coffey, R.J. Structural studies of human Naked2: A biologically active intrinsically unstructured protein. *Biochem. Biophys. Res. Commun.* **2006**, *350*, 911–915. <https://doi.org/10.1016/j.bbrc.2006.09.121>.
46. Peng, X.; Luo, Z.; Kang, Q.; Deng, D.; Wang, Q.; Peng, H.; Wang, S.; Wei, Z. FOXQ1 mediates the crosstalk between TGF- β and Wnt signaling pathways in the progression of colorectal cancer. *Cancer Biol. Ther.* **2015**, *16*, 1099–1109. <https://doi.org/10.1080/15384047.2015.1047568>.
47. Zhang, J.; Liu, Y.; Zhang, J.; Cui, X.; Li, G.; Wang, J.; Ren, H.; Zhang, Y. FOXQ1 promotes gastric cancer metastasis through upregulation of Snail. *Oncol. Rep.* **2016**, *35*, 3607–3613. <https://doi.org/10.3892/or.2016.4736>.
48. Li, Y.; Zhang, Y.; Yao, Z.; Li, S.; Yin, Z.; Xu, M. Forkhead box Q1: A key player in the pathogenesis of tumors (Review). *Int. J. Oncol.* **2016**, *49*, 51–58. <https://doi.org/10.3892/ijo.2016.3517>.

49. Kaneda, H.; Arao, T.; Tanaka, K.; Tamura, D.; Aomatsu, K.; Kudo, K.; Sakai, K.; De Velasco, M.A.; Matsumoto, K.; Fujita, Y.; et al. FOXQ1 is overexpressed in colorectal cancer and enhances tumorigenicity and tumor growth. *Cancer Res.* **2010**, *70*, 2053–2063. <https://doi.org/10.1158/0008-5472.Can-09-2161>.
50. Liu, J.Y.; Wu, X.Y.; Wu, G.N.; Liu, F.K.; Yao, X.Q. FOXQ1 promotes cancer metastasis by PI3K/AKT signaling regulation in colorectal carcinoma. *Am. J. Transl. Res.* **2017**, *9*, 2207–2218.
51. Nagy, A.; Munkacsy, G.; Gyorffy, B. Pancancer survival analysis of cancer hallmark genes. *Sci. Rep.* **2021**, *11*, 6047. <https://doi.org/10.1038/s41598-021-84787-5>.
52. Li, J.H.; Liu, S.; Zhou, H.; Qu, L.H.; Yang, J.H. starBase v2.0: Decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. *Nucleic Acids Res.* **2014**, *42*, D92–D97. <https://doi.org/10.1093/nar/gkt1248>.
53. Medico, E.; Russo, M.; Picco, G.; Cancelliere, C.; Valtorta, E.; Corti, G.; Buscarino, M.; Isella, C.; Lamba, S.; Martinoglio, B.; et al. The molecular landscape of colorectal cancer cell lines unveils clinically actionable kinase targets. *Nat. Commun.* **2015**, *6*, 7002. <https://doi.org/10.1038/ncomms8002>.
54. Wu, Y.F.; Wang, C.Y.; Tang, W.C.; Lee, Y.C.; Ta, H.D.K.; Lin, L.C.; Pan, S.R.; Ni, Y.C.; Anuraga, G.; Lee, K.H. Expression Profile and Prognostic Value of Wnt Signaling Pathway Molecules in Colorectal Cancer. *Biomedicines* **2021**, *9*, 1331. <https://doi.org/10.3390/biomedicines9101331>.
55. Lawler, T.; Parlato, L.; Warren Andersen, S. The histological and molecular characteristics of early-onset colorectal cancer: A systematic review and meta-analysis. *Front. Oncol.* **2024**, *14*, 1349572. <https://doi.org/10.3389/fonc.2024.1349572>.
56. Tang, J.; Peng, W.; Tian, C.; Zhang, Y.; Ji, D.; Wang, L.; Jin, K.; Wang, F.; Shao, Y.; Wang, X.; et al. Molecular characteristics of early-onset compared with late-onset colorectal cancer: A case controlled study. *Int. J. Surg.* **2024**, *110*, 4559–4570. <https://doi.org/10.1097/JS9.0000000000001584>.
57. Tashkandi, E. The Rise of Early-Onset Colorectal Cancer: Clinical Characteristics and Outcomes in a Young Patient Population. *Cancer Manag. Res.* **2025**, *17*, 2417–2427. <https://doi.org/10.2147/CMAR.S554044>.
58. Fink, S.P.; Myeroff, L.L.; Kariv, R.; Platzer, P.; Xin, B.; Mikkola, D.; Lawrence, E.; Morris, N.; Nosrati, A.; Willson, J.K.; et al. Induction of KIAA1199/CEMIP is associated with colon cancer phenotype and poor patient survival. *Oncotarget* **2015**, *6*, 30500–30515. <https://doi.org/10.18632/oncotarget.5921>.
59. Hua, Q.; Zhang, B.; Xu, G.; Wang, L.; Wang, H.; Lin, Z.; Yu, D.; Ren, J.; Zhang, D.; Zhao, L.; et al. CEMIP, a novel adaptor protein of OGT, promotes colorectal cancer metastasis through glutamine metabolic reprogramming via reciprocal regulation of beta-catenin. *Oncogene* **2021**, *40*, 6443–6455. <https://doi.org/10.1038/s41388-021-02023-w>.
60. Zhang, B.; Li, J.; Hua, Q.; Wang, H.; Xu, G.; Chen, J.; Zhu, Y.; Li, R.; Liang, Q.; Wang, L.; et al. Tumor CEMIP drives immune evasion of colorectal cancer via MHC-I internalization and degradation. *J. Immunother. Cancer* **2023**, *11*, e005592. <https://doi.org/10.1136/jitc-2022-005592>.
61. Liang, G.; Fang, X.; Yang, Y.; Song, Y. Silencing of CEMIP suppresses Wnt/beta-catenin/Snail signaling transduction and inhibits EMT program of colorectal cancer cells. *Acta Histochem.* **2018**, *120*, 56–63. <https://doi.org/10.1016/j.acthis.2017.11.002>.
62. Xu, G.; Zhao, L.; Hua, Q.; Wang, L.; Liu, H.; Lin, Z.; Jin, M.; Wang, J.; Zhou, P.; Yang, K.; et al. CEMIP, acting as a scaffold protein for bridging GRAF1 and MIB1, promotes colorectal cancer metastasis via activating CDC42/MAPK pathway. *Cell Death Dis.* **2023**, *14*, 167. <https://doi.org/10.1038/s41419-023-05644-z>.
63. Zhang, L.; Yang, F.; Zhang, X.; Dong, D.; Fang, X.; Meng, L.; Chang, T. Identification of the Pro-Tumorigenic Role of the ELK4-METTL3-CEMIP Axis in Colorectal Carcinoma: Promotion of Cancer Cell Stemness and Malignant Phenotypes. *J. Gastroenterol. Hepatol.* **2025**, *40*, 1963–1975. <https://doi.org/10.1111/jgh.17030>.
64. Gong, Z.; Feng, Y.; Tu, J. Prediction of immunotherapeutic responses by a classifier model based on inflammation-associated tumor microenvironment signatures in colorectal cancer. *Discov. Oncol.* **2026**, *17*, 378. <https://doi.org/10.1007/s12672-026-04548-6>.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.