CCDC12 Promotes Tumor Development and Invasion Through the Snail Pathway in Colon Adenocarcinoma

Fengying Du
Shandong Provincial Hospital
Leping Li
Shandong Provincial Hospital
Qiang Wang
Shandong Provincial Hospital
Wenting Pei
Qilu Children's hosptial of Shandong University
Hongqing Zhuo
Shandong Provincial Hospital
Lipan Peng
Shandong Provincial Hospital
Tao Xu
Shandong Provincial Hospital
Changqing Jing
Shandong Provincial Hospital
Jizhun Zhang (zhangjizhun2007@163.com)
Shandong Provincial Hospital https://orcid.org/0000-0001-7383-9577

Research

Keywords: CCDC12, EMT, Snail pathway

Posted Date: August 14th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-57291/v1

Version of Record: A version of this preprint was published at Cell Death & Disease on February 1st, 2022. See the published version at https://doi.org/10.1038/s41419-022-04617-y.

Abstract

Background: To determine the role of Coiled-coil domain containing 12 (CCDC12) in colon adenocarcinoma (COAD).

Methods: We used integrative expression Quantitative Trait Loci (eQTL) analysis to identify risk single nucleotide polymorphisms (SNPs) and associated genes. Immunohistochemical staining (IHC) and western blotting (WB) were used to detect *CCDC12* expression. We selected SW480 and LOVO cell lines to knock down *CCDC12*, while HCT116 and SW480-KD cell lines were up-regulated. Then performed functional studies and established a xenograft tumor model in BALB/c mice. Four plex Isobaric Tags for Relative and Absolute Quantitation (iTRAQ) assays were performed to determine its function and potential regulatory mechanism. Overexpressed Snail and knocked down *CCDC12* subsequently in SW480 cells to find them association.

Results: Integrative eQTL analysis found that rs8180040 was significantly associated with *CCDC12* in COAD patients. IHC and WB confirmed *CCDC12* was highly expressed in COAD tissues (IHC 51/75 vs 8/75, WB 10/12 vs 2/12). This was consistent with RNA-Seq data from TCGA database (*P* value < 0.001). Knockdown of *CCDC12* could significantly reduce proliferation, migration, invasion and tumorigenicity of colon cancer cells, while exogenous overexpression of *CCDC12* had the opposite effect. iTRAQ assays demonstrated that overexpression of *CCDC12* would change proteins on adherens junction pathway. Overexpression of Snail did not significantly change *CCDC12* levels in SW480 cells, while knock down of *CCDC12* reduced that of Snail.

Conclusions: *CCDC12* plays a significant role in tumorigenesis, development and invasion of COAD and may affect the epithelial to mesenchymal transformation process of colon cancer cells by regulating the Snail pathway.

1. Background

The incidence of colon cancer is the fourth highest for malignant tumors, with over one million new colon cancer patients worldwide each year(1). In the Asian population, approximately 90% of colon cancers are histologically classified adenocarcinomas, and the prognosis of these patients is poor (2, 3). Refractory and metastatic colon adenocarcinoma (COAD) has been a major problem worldwide (4-6). Somatic mutations and activation of key oncogenic pathways have often been observed in COAD. Hence, it is essential to understanding the mechanism of refractory and metastatic COAD. In addition, identifying effective prognostic biomarkers for high risk patients with COAD is vital.

Coiled-coil domain containing 12 (*CCDC12*) is an evolutionarily conserved protein that encodes a coiled-coil domain. It is located in the 3p21.31 region of human chromosome 3(7). *CCDC12* has been reported to be associated with erythroid differentiation(8), and split-ubiquitin system(9). In addition, coiled-coil domain containing family members have been associated with tumor cell proliferation. *CCDC106* is associated with the progression and poor prognosis of non-small cell lung cancer(10), while *CCDC67* has

been demonstrated to inhibit the proliferation of papillary thyroid carcinomas(11). A genome-wide association study (GWAS) in China identified colorectal cancer risk single nucleotide polymorphism (SNP) rs1076394 as an expression Quantitative Trait Loci (eQTL) for *CCDC12*(12). However, the specific carcinogenesis of *CCDC12* has not been deciphered.

In this study, we demonstrated that high-expression levels of *CCDC12* in COAD was closely associated with tumor development and aggressive. *CCDC12* promoted COAD tumor cell proliferation, invasion, migration, and inhibited apoptosis in *in vitro* and *in vivo* experiments. Furthermore, *CCDC12* could regulate epithelial to mesenchymal transformation (EMT) of COAD cells through the Snail pathway. Our study demonstrated a biological link between *CCDC12* and COAD, which could be used as a potential therapeutic target.

2. Methods

2.1 Colorectal cancer (CRC)-associated SNPs and germline genotype data

CRC-associated SNPs were extracted from the National Human Genome Research Institute (NHGRI) GWAS database including 50 CRC risk loci (Supplement to table S1). The datasets for germline genotypes, ancestry, expression profiles, methylation, somatic copy number aberrations, germline copy number aberrations regarding CRC were downloaded from The Cancer Genome Atlas (TCGA) portal. SNP loci with minor allele frequency (MAF) > 0.05 from TCGA (subjects) and HapMap cell lines (controls) were downloaded on EIGENSTRAT and the top two principal components were retrieved (Figure S1). We calculated the average of segmented copy-number scores of genetic interval between the transcription start and end sites as gene-based somatic copy-number measure (Figure S2A). CpG methylation status was determined by discretization CpG methylation value with cut-off values of 0.2, 0.4, 0.6, 0.8, and 1.0 (Figure S2B). Afterward, we calculated the expression levels for each gene as TPM values.

2.2 Association analysis and eQTL analysis

eQTL analysis was performed according to the flowchart described in Li et al. (13). Briefly, the expression data of gene was adjusted for somatic copy-number effects and CpG methylation status using a multivariate linear model. The *P value* corresponds to the regression coefficient based on residual expression levels and germline genotype. We performed cis-eQTL analysis between 656 857 SNP loci and corresponding mRNA transcripts. Then excluded 243 359 SNP loci with MAF < 0.05 and their genes with absent calls >90% and false discovery rate (FDR) > 0.1. With the 50 SNPs from NHGRI GWAS database, 18 SNPs were present in TCGA germline genotype. For variants not directly genotyped, we used proxy SNPs (nearest SNP with linkage disequilibrium > 0.5). SNAPinfo software was selected to obtain pairwise linkage disequilibrium between SNPs(13). In cis-eQTL analysis, we evaluated the association between genotype of given SNP locus and the transcripts located within ± 1Mb regions. For risk SNP locus, the

relationship between it and transcripts at a genome-wide level was evaluated. For each target gene, 1-50 Kb regions on either side of transcription start site were considered putative enhancer regions. They were overlapped using ENCODE DNaseI hypersensitivity data from HCT116 cell line and then analyzed for transcription factor (TF) DNA binding motif enrichment. Hypergeometric distribution test was used for overlap analyses with a significance level of *P value* < 0.05. TFs that satisfied the above criteria were considered as candidates for trans-acting risk SNPs.

2.3 Ancestry verification samples, tissue samples and TMAs

130 cases of CRC patients from Shandong Provincial Hospital were used for ancestry verification. Germline genotypes were measured with patients' peripheral blood samples and matched tumor samples. Another 24 fresh tissues were obtained from the Department of Gastrointestinal Surgery of Shandong Provincial Hospital, including 12 colon adenocarcinomas and 12 paired adjacent normal colons. All patients had not received preoperative radiotherapy or chemotherapy. All procedures were performed in accordance with the International Ethical Guidelines for Biomedical Research Involving Human Subjects (CIOMS) and Declaration of Helsinki. We obtained written informed consent from all patients. Tissue microarrays (TMA) purchased from Molbase Co. Ltd.(Shanghai, China), consisted of 75 pairs of human COAD and their adjacent colon tissue.

2.4 Immunohistochemistry staining

Immunohistochemical staining was performed using Power-Vision two-step tissue staining kit (ZSGB-BIO, Cat. PV-6001, Beijing, China). After deparaffinization and rehydration, tissue slides were incubated with 3% H_2O_2 for 10 minutes at room temperature. 10 mmol/L EDTA solution used for antigen retrieval. Primary antibodies were incubated overnight at 4℃ and then secondary antibodies incubated at 37℃ for 30 minutes. After washed, stained with DAB-H_2O_2 and counterstained with hematoxylin. The results of IHC were evaluated using H-scores by 3 researchers independently (Table S2).

2.5 Cell lines and cell culture

The human colon cancer cells selected were HCT116, HT29, LOVO, SW480 and SW620, CCD-18Co cell line as the control. All cell lines were obtained from the American Type Culture Collection (ATCC, USA). HCT116, HT29 and CCD-18Co cell lines were cultured in DMEM/high glucose medium (Hyclone, Cat.SH30022.01B, USA), LOVO cell line was F-12K medium (Gibco, Cat.21127022, USA), SW480 and SW620 cell lines were L-15 medium (Gibco, Cat.11415064, USA). Each medium was supplemented with 10% fetal bovine serum (Hyclone, Cat.SH30087.01, USA) and 1% penicillin-streptomycin (Hyclone, Cat.SH30010, USA). Cells were cultured at 37℃ containing 5% CO_2.

2.6 RNA interference and overexpression of *CCDC12*

Three *CCDC12* short interfering RNA (siRNA) were purchased from RiboBio (Cat. 140901180506, Shanghai, China) and were transfected into SW480 and LOVO cell lines using Lipofectamine™ RNAiMAX reagent (Invitrogen, Cat.13778075, USA) for 24 hours at 37 ℃. Lentivirus with puromycin resistance expressing green fluorescent protein was used to overexpress *CCDC12* in HCT116 cells and interfered SW480 (SW480-KD, transfected siCCDC12), designed and synthesized by Genechem (Shanghai, China). 72 hours after transfection, cells were exposed to puromycin for 48 hours to select.

2.7 Over expression of Snail in SW480 cell lines

Lentivirus expressing *SNAI1* was used to infect SW480 cell line and expressed red fluorescent protein. Seventy-two hours later, fluorescence of red fluorescent protein was observed under a fluorescent microscope (Olympus, IX71, Japan) then they were used for subsequent experiments.

2.8 Western blotting

The extracted proteins from cells and tissues were electrophoresed on a 10% SDS-PAGE, and then transferred onto a 0.45μm Immobilon-P Transfer Membrane (Millipore, Cat.IPVH00010, USA) using the wet transfer method. Incubated with primary antibody overnight at 4℃ and then incubated with corresponding secondary antibody for 1 hour at room temperature. Bands were visualized using ECL kit (Millipore, Cat.WBKLS0500, USA) and the Amersham Imager 680 system. Primary antibodies used for western blotting are listed in Table S2.

2.9 Real-time Quantitative Polymerase Chain Reaction (RT-PCR)

Total RNA was extracted with RNA isolation kit (Invitrogen, USA) and reverse transcribed to cDNA using the Reverse Transcription System (Promega, USA). Quantitative RT-PCR was performed with SYBR Green qPCR SuperMix (Invitrogen, USA) and ABI PRISM® 7500 Sequence Detection System based on the manufacturer's instructions. 18srRNA was used as the internal reference control. Primer sequences were as follows: 18srRNA forward 5'-CCTGGATACCGCAGCTAGGA-3', reverse 5'-GCGGCGCAATACGAATGCCCC-3', CCDC12 forward 5'-CTGACTGGGACCTCAAGAGA-3', reverse 5'-CCTTTCAGCCTTTCACGGAT-3', Snail forward 5'-GAGGCGGTGGCAGACTAGAGT-3', reverse 5'-CGGGCCCCCAGAATAGTTC-3'.

2.10 Colony-forming and MTS assays

100 cells in logarithmic growth phase were resuspended in 300μl medium and then seeded into a 6-well plate. After colony formation, stained with 1% crystal violet solution for 20 minutes. Number of colonies were counted under a microscope.

In MTS assay, 1×10^4 cells were seeded into a 96-well plate. CellTiter 96® A Q_{ueous} One Solution Cell Proliferation Assay (Promega□Cat.G3582, USA) was used to measure cell proliferation and was

performed based on the manufacturer's instructions. OD was measured using an Multiscan MK3 microplate reader at 490nm.

2.11 Wound-healing assays

Cells were seeded into a 6-well plate and cultured until 95% confluent. Monolayer cell was scraped off using a pipette tip in the middle of plate. Cell migration was measured every 6 hours using the Image Pro-Plus 6.0 and migration rate was calculated with ($Distance_{0h}$ - $Distance_{different\ time\ points}$) / $Distance_{0h}$.

2.12 Cell invasion assays

Transwell chambers (BD, Cat.353097) with Matrigel (BD, Cat.356234, USA) were used for invasion assays. 100μl of cells suspension (1×10^5 cells) with serum-free medium were placed in the upper chamber. The bottom chamber contained medium with 20% serum. Cells were incubated for 24 hours at 37℃, 4% paraformaldehyde was used to fix the cells for 15 minutes, and then stained with crystal violet solution. The cells passing through the chamber were observed with a microscope.

2.13 Apoptosis assays

The Annexin V-FITC apoptosis detection kit (Keygen, Cat.KGA106, Jiangsu, China) was used per manufacturer's instructions. 1.25μl Annexin V-FITC reagent was added to 500μl of cell suspension (1×10^6/ml), and then incubated for 15 minutes at room temperature in the dark. After centrifuged at 1 000×g for 5 minutes, supernatant was removed, and cells were resuspended in 0.5 ml pre-cooled binding-buffer. Then, 10 μl Propidium Iodide was added and incubated in the dark before being read on the BD FACSCalibur CellSorting System.

2.14 Cell cycle analysis

Cell Cycle Detection Kit (Keygen, Cat.KGA511, Jiangsu, China) was used for cell cycle analysis. 5 μl (10mg/ml) of RNase A was added to cells and incubated at 37℃ for 1 hour. Afterward, 50 μg/ml PI and 0.2% Triton X-100 were added and incubated at 4℃ in the dark for 30 minutes. BD FACSCalibur CellSorting System was used to measure cell cycle phases. $2\text{-}3\times10^4$ cells were counted and analyzed using ModFit software.

2.15 Xenograft mouse models

4-week-old BALB/c nude mice were purchased from Charles River Laboratories (Beijing, China) fed on ordinary diet. Xenograft tumors were established by subcutaneous injection of 200μl cell suspension (5×10^5 cells) into the underarms or backs of nude mice. The tumor volume was calculated using the

following formula, Tumor Volume(mm^3) = (Long diameter × Short diameter^2)/2. Mice were euthanized 30 days after inoculation, and tumors removed for subsequent analysis.

2.16 4 plex Isobaric Tags for Relative and Absolute Quantitation (iTRAQ) assays

Proteins was extracted from *CCDC12* over-expressing HCT116 cells and control HCT116 cells. The Bradford quantitative method was used to determine the total protein content. Proteins were reductively alkylated using DTT and TEAB. After enzyme digestion, proteins were acidified with 0.1% FA. The components were grouped based on high pH C18 (chromatograph was Thermo DINOEX Ultimate 3000 BioRS; analytical column was Durashell C18, 5μm, 100 Å, 4.6×250mm) and the isolated components were analyzed by LC-MS/MS (Thermo Fisher Q-exactive HF-X; AB SCIEX analytical column: 75 μm inner diameter, packed with 3 μm, 120 Å ChromXP C18 column, 10 cm long; eksigent Chromxp Trap Column: 3 μm C18-CL, 120 Å, 350 μm×0.5 mm). The results were analyzed through Cluster of Orthologous Groups of proteins (COG) analysis, Gene Ontology (GO) analysis, and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment.

2.17 Statistical Analysis

All statistical analysis was performed using R 3.5.1. Data were expressed as mean ± SD. One-way analysis of variance and Student's T test were used to analyzed differences among groups. χ^2 test or linear correlation was used to determine the correlation between *CCDC12* expression and clinicopathological features. Kaplan-Meier method was used to generate survival curves with log-rank test. MAF > 0.05, FDR < 0.1, α = 0.05 and *Pvalue* < 0.05 with two sides were considered statistically significant.

3. Results

3.1 SNP rs8180040 is significantly associated with *CCDC12* expression based on integrative eQTL analysis

Ancestry verification confirmed that the 130 samples in our study was HapMap CEU, and table S3 summarizes their clinical and pathologic data. Of all the risk SNP loci, only rs10505477 and rs6983267 were in high pairwise linkage disequilibrium (R^2 = 0.875). After filtering out FDR < 0.1, we obtained 5 029 762 SNP-genes and identified 2 030 significant associations with *P value* < 4.03×10^{-5}. They mapped to a total of 1 964 SNP loci and 478 unique target genes (Table S4). Of the target genes, 332 (69.5%) were found to be regulated by a single cis-acting SNP locus. After adjusting for correlated loci by stepwise feature selection, 254 genes (53.1%) could be explained by multiple SNP loci (median = 4), which suggested the existence of multiple independent eQTLs. Of the 1,964 cis-acting eQTL loci, 1 899 were associated with one target gene, and suggested that the associations were highly locus specific.

Transcript levels of 6 638 genes (55.6%) were significantly affected by the somatic copy-number changes in the corresponding coding regions (FDR < 0.1). In addition, among the 478 target genes of the cis-acting SNP loci, 299 (62.6%) were significantly associated with somatic copy number. We identified 6 662 transcripts (55.8%) that were affected by CpG island methylation in the promoter region (FDR < 0.1). Among these genes, 302 (63.2%) were also target genes of eQTLs (Figure 1A). Subsequently, we identified 25 SNP-gene expression associations in 22 loci with nominal *P value* < 0.05 for genes within 2Mb of the risk SNPs (Table S5). After correcting for multiple-testing (FDR < 0.05), only one SNP-gene association was found to be significant (SNP rs8180040 with genes *CCDC12*, Table S5 and Figure 1A).

3.2 High expression of *CCDC12* in COAD is associated with poor prognosis

Our previous study found that *CCDC12* may be associated with colon cancer and was a potential proto-oncogene(14). To determine the expression levels of *CCDC12* in COAD, we performed immunohistochemistry (IHC) with TMAs that contained 75 pairs of COAD and adjacent normal tissues. Positive staining rate of *CCDC12* in COAD tissues were significantly higher compared to normal colon tissues (51/75 vs 8/75, *P value* < 0.001, Figure 1B). In addition, we performed western blotting with fresh tumor and adjacent normal tissues from 12 patients. The results showed that the expression levels of *CCDC12* in 10 colon cancer tissue samples were significantly higher compared to normal adjacent tissues (Figure 1C). We then analyzed RNA-Seq data from TCGA database, which included 456 COAD patients and 41 corresponding normal tissues. It found that *CCDC12* was statistically significantly overexpressed in COAD (*P value* < 0.001, Figure 1D). Subsequently, we determined the correlation between expression level of *CCDC12* and prognosis of patients with COAD. Our analysis demonstrated that higher expression of *CCDC12* was correlated with poor prognosis of COAD patients grouped by gender (*P value* = 0.0042, Figure 1E).

3.3 Knockdown of *CCDC12* inhibits proliferation, migration, invasion and promoted apoptosis.

To investigate the biological role of *CCDC12* in colon cancer, we first measured the expression levels of *CCDC12* in five different colon cancer cell lines by qRT-PCR, and the CCD-18Co as the control (Figure S3). The expression levels of *CCDC12* in LOVO and SW480 cell lines were relatively high, hence we reduced their expression levels using si-CCDC12 RNA. qRT-PCR results demonstrated that siRNA1 had the best knockdown efficiency for *CCDC12* in both SW480 and LOVO cell lines (Figure S4). So, we used siRNA1 for all subsequent functional assays. *In vitro* experiments were performed with *CCDC12* knockdown group (CCDC12-KD, cells transfected with siRNA1), normal group (NC, un-transfected cells) and blank control group (Vec, cells transfected with empty vector).

Colony-forming assays demonstrated that both SW480 and LOVO cell lines in CCDC12-KD group had reduced ability to form colonies compared to cells in the NC and Vec groups (Figure 2A). Cell cycle assay

demonstrated that compared to cells in others group, cells in the CCDC12-KD group were predominately in the G0/G1 phase, with a reduced proportion of cells in the S phase. In addition, it was observed that in SW480 cell lines, the ratio of cells in the G2/M phase were decreased (Figure 2B). Using MTS assays, we observed that the rates of cell proliferation in the CCDC12-KD group were significantly lower on the third day compared to that in the NC and Vec groups (Figure S5A). In wound-healing assays, we observed slight differences in migration at 6hrs between cells in the CCDC12-KD group and NC group. At 24hrs, cell migration in the CCDC12-KD group were significantly reduced (Figure 2C). Matrigel-coated transwell chambers were then used to evaluate differences in cell invasion between the three treatment groups. We found that the cell invasion ability in the CCDC12-KD group were much lower, especially in the SW480 cell line (Figure 2D). Using Annexin V-FITC to measure apoptosis rates, we observed that the total apoptosis rate (UR+LR) of cells in the CCDC12-KD group were increased. There were no differences in early apoptosis between three groups, however there were differences in late apoptosis (Figure 2E).We then injected CCDC12-KD SW480 or NC SW480 cells subcutaneously into the back of BALB/c nude mice to generate xenograft models. After 30 days, we observed that the volume and weight of tumors in mice in CCDC12-KD group were much smaller (Figure 2F). All these results indicated that knockdown of *CCDC12* could effectively inhibit colon cancer cell proliferation, invasion, migration, and promoted apoptosis *in vivo* and *in vitro*.

3.4 Overexpression of *CCDC12* promotes proliferation, migration, invasion, and inhibits apoptosis

Based on expression levels of *CCDC12* in colon cancer cell lines (Figure S3), we selected HCT116 cell line and *CCDC12*-knocked SW480 cell line to overexpress *CCDC12*. In addition, we overexpressed *CCDC12* stably (OE-CCDC12 group) in these two cell lines (Figure S6) and set normal group (NC) and blank control group (Vec) as described in the last section.

In colony formation assays, cells in the OE-CCDC12, NC, and Vec groups formed colonies on the seventh day of culture. However, cells in the OE-CCDC12 group generated a higher number of colonies compared to the other two groups (Figure 3A). In addition, we observed a higher proportion of OE-CCDC12 HCT116 cells in the G0/G1 phase, with a lower proportion of cells in the S and G2/M phase. For SW480-KD cells in the OE-CCDC12 group, cells were more arrested in the G0/G1 phase, while reduced in the G2/M phase (Figure 3B). We found a higher statistically significant rate of cell proliferation in OE-CCDC12 group through MTS assays (Figure S5B). Next, we evaluated cell migration ability using wound-healing assays. There were slight differences in migration ability between cells in OE-CCDC12 group and NC group at 24 hours. However, after 48 hours, the cell migration ability in OE-CCDC12 group was much higher (Figure 3C). We then measured the invasive ability of cells using matrigel-coated transwell chambers. As shown in Figure 3D, the invasive ability of cells in OE-CCDC12 group was higher compared to cells in NC and Vec groups, and especially for SW480-KD cell lines. Furthermore, we observed that the total apoptotic rate of cells was reduced by approximately 50% in OE-CCDC12 group (Figure 3E). This was more evident in HCT116 cell lines, specifically, the ratio of cells in early apoptosis was reduced.

We next performed *in vivo* studies by injecting OE-CCDC12 HCT116 or NC HCT116 cells subcutaneously into the axilla of nude mice. Mice were then euthanized 4 weeks later to determine tumor growth. Tumor volume and weight in nude mice injected with OE-CCDC12 HCT116 cells were larger compared to NC group (Figure 3G). These findings demonstrate that over-expression of *CCDC12* increases colon cancer cell proliferation, migration, and invasiveness both *in vivo* and *in vitro* while reducing apoptosis levels.

3.5 *CCDC12* induces epithelial-mesenchymal transition (EMT) to aggravate COAD

iTRAQ was used to quantitatively determine differential protein expression in cells after *CCDC12* overexpression, and we compared their expression in OE-CCDC12 HCT116 cells and NC HCT116 cells. We set the threshold for fold change (FC) to ≥ 1.3 or ≤ 0.77 and q value < 0.05. 127 proteins had higher expression in OE-CCDC12 HCT116 cells, while 42 proteins had lower expression (Figure 4A). We analyzed the differentially expressed proteins using the COG database to predict possible functions and functionally classify these proteins (Figure 4B). Enriched protein classifications included translation, ribosomal structure, and biogenesis; transcription; replication, recombination, and repair; posttranslational modification, protein turnover, chaperones; signal transduction mechanisms, etc., all of which are important functional classifications associated with cancer. We then performed cluster analysis on differentially expressed proteins. The top 50 differentially expressed proteins clustered into several groups (Figure 4C) to classify *CCDC12* interacting proteins, and the full heatmap shown in Figure S7. GO analysis was used to further classify the function of these differentially expressed proteins. We have separately annotated the increased and decreased proteins, and we only describe the increased proteins. As for Biological Process (BP) classifications, proteins were mostly annotated into transcription (DNA-templated), negative regulation of transcription from RNA polymerase promoter, cell division and cell proliferation (Figure 4D); for Molecular Function (MF), the majority were protein binding, enzyme binding, protein kinase binding and Zinc ion binding (Figure 4E); for Cellular Component (CC), cytoplasmic ribonucleoprotein granule and nucleus (Figure 4F). In terms of signaling pathways, adherens junction, viral carcinogenesis, and epithelial cell signaling in Helicobacter pylori infection were annotated (Figure 4G). These results suggested that *CCDC12* may regulate EMT in COAD, more specifically via the regulation of intercellular adherens junction. To support these findings, we performed western blotting using HCT116 cells overexpressing *CCDC12* and compared them with naive HCT116 cells. As shown in Figure 4H, *CCDC12* expression was associated with the biomarkers in EMT process. Overexpression of *CCDC12* induced significant changes in E-cadherin, Vimentin, Fibronectin, Matrix Metallopeptidase 9 (MMP 9), Snail, and Slug protein levels. In particularly, the expression levels of Snail and Slug dramatically changed in HCT116 cells after overexpressing *CCDC12.* In summary, *CCDC12* regulated COAD by altering the expression levels of several biologically functional proteins that were associated with EMT, especially zinc finger transcription factors.

3.6 *CCDC12* regulates EMT in COAD via zinc finger transcription factors

Using red fluorescent protein-labeled lentivirus, we overexpressed SNAI1 in SW480 cells (Figure S8A). After determining the rate of transfection by a fluorescence microscope, we measured expression levels of Snail by western blotting (Figure S8B) and qPCR (Figure S8C). We then used siRNA1 to knockdown *CCDC12* in SW480 cells, where overexpressed Snail. Western blotting demonstrated that compared to cells in NC group, expression levels of *CCDC12* in OE-Snail cells did not change. Knockdown of *CCDC12* resulted in the expression level of Snail decreasing to a certain extent (Figure 5A). To determine whether *CCDC12* regulates EMT dependent on Snail, we used transwell chamber to evaluate the invasive ability of cells knocked down *CCDC12* after overexpressing Snail. Overexpression of Snail significantly increased cell invasion. However, knocking down *CCDC12* in Snail overexpressing cells reduced cell invasion but was still higher compared to cells in the NC group (Figure 5B). These results suggested that *CCDC12* regulates EMT in COAD by affecting Snail expression.

4. Discussion

Adenocarcinoma is the main pathological type of colon cancer and accounts for more than 95%, especially in Asia. Based on its biological characteristics, COAD is highly metastatic, which results in poor five-year patient survival (1-3). Patients who have undergone systemic chemotherapy after surgery have not shown improvements in survival rates(15). With the development and innovation of surgical techniques, the prognosis of patients with colon cancer seems to have improved. However, these new surgical technologies have not been fully adopted(16, 17). Concurrently, large clinical trials using targeted therapies have been performed worldwide(4, 18, 19), to find treatments that could effectively improve patient prognosis and inhibit colon cancer metastasis.

The *CCDC12* gene, located on 3p21.31 of chromosome 3, has been reported to play a role in colorectal tumorigenesis(12). *CCDC12* has been demonstrated to accelerate the growth of K562 cells by up-regulating CD235, ε-globin, and γ-globin in human chronic myeloid leukemia(8). Anne et. al. reported that *CCDC12* may be associated with ubiquitination(9), which is particularly critical in tumor cells and could participate in the modification and degradation of some cancer factors to affect the biological behavior of tumors. In addition, Ke et al. using GWAS found that *CCDC12* may be a potential risk gene for colorectal cancer and associated with a potential regulatory variant, rs1076394(12). In this study, we demonstrated that *CCDC12* was highly expressed in colon adenocarcinomas and may affect cancer metastasis by regulating EMT.

We analyzed GWAS data from NHGRI database, and then performed deep mining of colorectal cancer-related data from TCGA. Our team found numerous colorectal cancer-related risk SNPs, which were verified through 130 additional samples. With this, we identified rs10505477 and rs6983267 that had high pairwise linkage disequilibrium. After integrative eQTL-based analysis, we identified 25 SNP-gene expression associations in 22 risk SNP loci, and multiple corrections were performed to improve the

reliability. We found SNP rs8180040 and *CCDC12* had a significant correlation, both of which were located on chromosome 3 and had a strong geometric correlation. SNP rs8180040 was located at chr3: 47347457 (GRCh38.p12) with T > A allele (T = 0.5992 and A = 0.4008). The gene frequency of rs8180040 was slightly different in different populations. The Asian population had the highest allele A frequency of 0.5, especially in East Asian populations; the African population had the lowest frequency of 0.23; the European population was close to the average (A = 0.4088), (12, 20, 21) and our results were similar to previous studies. Hence, we inferred that rs8180040 may affect the occurrence of colorectal cancer by regulating *CCDC12* expression.

We then measured the expression levels of *CCDC12* in cancer and normal tissues by IHC staining. Using 75 pairs of COAD and matched adjacent normal mucosa samples demonstrated that *CCDC12* was frequently overexpressed in COAD (68.0%). We sequentially performed confirmatory western blotting through 12 fresh COAD samples and demonstrated higher expression levels of *CCDC12* compared to matched adjacent normal mucosa samples. Based on a large-scale RNA-Seq dataset in TCGA, we confirmed that *CCDC12* was frequently overexpressed in COAD tissues. Cox regression analysis performed by gender showed that *CCDC12* was a prognostic factor. Interestingly, in female patients, *CCDC12* overexpression leads to a poorer prognosis, but not in male patients. This was the focus of our study in subsequent experiments. These findings demonstrated that *CCDC12* plays a potential role in COAD tumorigenesis and development.

We next performed a series of *in vivo* and *in vitro* experiments to determine how *CCDC12* regulates the proliferation, migration, and invasion of COAD cells to understand its biological role. After knockdown of *CCDC12*, cell proliferation, invasion, and migration were reduced, while rate of apoptosis increased. In addition, knockdown of *CCDC12* resulted in a higher proportion of cells in the G0/G1 phase. In contrast, overexpression of *CCDC12* increased cell proliferation, invasion, and migration, while reducing apoptosis levels. Interestingly, overexpression of *CCDC12* still resulted cells in a block in G0/G1 phase. If cells do not enter S phase and are in a quiescent state, which meaning not participate in cell division. Additional studies are required to understand these findings in the future. Using xenograft tumor mouse models, knockdown of *CCDC12* significantly reduced tumor size and weight, while *CCDC12* overexpression promoted xenograft tumor growth. These findings strongly demonstrate that increased expression of *CCDC12* was essential for COAD tumorigenesis and development.

We next investigated the role of *CCDC12* in COAD metastasis. Through iTRAQ assays, we observed that overexpression of *CCDC12* could induce expression levels changes in 169 proteins. In COG analysis, about 1,000 proteins annotated with definite functions and enriched in the following annotations: translation, ribosomal structure and biogenesis (22); transcription(23); replication, recombination and repair(24); posttranslational modification, protein turnover, chaperones(25); and signal transduction mechanisms(26), all of which have been strongly associated with cancer. Cluster analysis divided the differentially expressed proteins into two groups and was in good agreement with the original groups. For proteins that changed in expression levels after *CCDC12* overexpression, we performed GO analysis to annotate their functions. For BP classification, the majority of proteins were annotated to transcription

(DNA-templated), negative regulation of transcription from RNA polymerase □ promoter, cell division, and cell proliferation. These constituted the core functions involved in cancer biology, which were based on changes in protein-coding genes and non-coding regulatory elements(27), especially for DNA template transcription. We also classified differentially expressed proteins into MF terms, which included zinc ion binding, protein binding, enzyme binding, and protein kinase binding. Metal ions form the basis for higher-order structures in protein assembly and provide short peptides chemical reactivity(28). In addition, zinc ions play a key role in homeostasis, immune function, and apoptosis(29), and may also induce p53 misfolding(30). With regards to CC classification, several proteins were annotated to the nucleus, cytoplasmic ribonucleoprotein granule, and nuclear chromatin. Chromatin state significantly affects the occurrence and treatment of cancers through specific regulatory mechanisms(31), additionally, the state of plasma membrane is critical for intercellular communication(32). What's more, we observed that *CCDC12* was localized in the nucleoplasm, which suggested that *CCDC12* may affect proteins in the nucleus, such as Snail (localized in the nucleus and cytosol). KEGG pathway enrichment analysis demonstrated that cancer-related pathways such as adherens junction and viral carcinogenesis were annotated. Early EMT is associated with the overall deterioration of cell-cell adhesion, which triggers front-rear polarization of cells required for migration(33). It is gratifying that in the adherens junction pathway, Snail and Slug are its typical representatives, both of which could be detected in the nucleus. And bioinformatics analysis based on the results of iTRAQ assay further confirmed that overexpression of *CCDC12* induced changes in expression levels of proteins, especially in the nucleus. All these findings demonstrated that *CCDC12* may regulate EMT of colon cancer cells through Snail located in the nucleus.

To confirm that *CCDC12* overexpression was involved in the invasion and malignancy of COAD, we measured typical EMT biomarkers by western blotting. Overexpression of *CCDC12* reduced E-cadherin level, but increased expression levels of Fibronectin, Vimentin, MMP 9, Snail, and Slug. During EMT, epithelial cells lose their cell polarity and connection with the basement membrane, resulting in the ability to migrate and invade, resist apoptosis, and degrade extracellular matrix(34), which were consistent with our *in vitro* experiments. iTRAQ assays and western blotting data suggested that Snail may be a downstream target of *CCDC12*, because not only Snail expression levels positively correlated with that of *CCDC12* but also showed the highest fold change in expression levels by western blotting (Figure 4H). Through lentivirus, we made Snail overexpressed in SW480 cells, and then knocked down *CCDC12*. Overexpression of Snail had no significant impact on *CCDC12* expression levels, while knockdown of *CCDC12* decreased Snail expression levels (Figure 5A). Additionally, we performed transwell assays and demonstrated that Snail overexpression effectively increased the invasion ability of colon cancer cells, and after subsequent knockdown of *CCDC12*, the invasive ability promoted by overexpression of Snail was corrected to some extent. These demonstrated that Snail was a downstream target of *CCDC12.* As to how *CCDC12* regulates Snail to affect colon cancer EMT requires additional follow-up studies.

5. Conclusion

We demonstrated that *CCDC12* plays an important role in the tumorigenesis and development of COAD. Higher expression levels of *CCDC12* in COAD was significantly associated with malignancy and

metastasis. *CCDC12* may be a potential oncogene that regulates EMT in COAD by affecting the Snail pathway, which and then affects the prognosis of patient. Additionally, *CCDC12* is expected to be a potential therapeutic target for COAD.

Abbreviations

Full name	Abbreviation
Coiled-coil domain containing 12	CCDC12
Colon adenocarcinoma	COAD
Expression Quantitative Trait Loci	eQTL
Single nucleotide polymorphism	SNP
Immunohistochemical staining	IHC staining
Western blotting	WB
Isobaric Tags for Relative and Absolute Quantitation	iTRAQ
Genome-wide association study	GWAS
National Human Genome Research Institute	NHGRI
The Cancer Genome Atlas	TCGA
Minor allele frequency	MAF
False discovery rate	FDR
Transcription factor	TF
The American Type Culture Collection	ATCC
Short interfering RNA	siRNA
Real-time quantitative polymerase chain reaction	RT-PCR
Cluster of Orthologous Groups of proteins	COG
Gene Ontology	GO
Kyoto Encyclopedia of Genes and Genomes	KEGG
Epithelial-mesenchymal transition	EMT
Fold change	FC
Biological Process	BP
Molecular Function	MF
Cellular Component	CC
Matrix Metallopeptidase 9	MMP9

Declarations

Ethics approval and consent to participate

Institutional review board approval was obtained prior to data abstraction and animal experiments by Ethics Committee of Shandong Provincial Hospital (NO.2017007) and individual patient written informed consent was also obtained. This trial was conducted in accordance with the Declaration of Helsinki.

Consent for publication

Not applicable

Availability of data and materials

All data generated or analysed during this study are included in this published article and its supplementary information files. The more data during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the National Natural Science Foundation of China (Grant No.81702363) in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Authors' contributions

LPL and CQJ were responsible for designing of the study and critical review of manuscript; JZZ and WTP were responsible for designing and performing of the study, literature research and manuscript writing; LPP, TX and HQZ were responsible for data collections; QW and JZZ were responsible for data analysis; FYD, WTP and HQZ performed experiments. All authors read and approved the final manuscript.

Acknowledgements

Not applicable

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA

Cancer J Clin. 2018;68(6):394-424.

2. Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, et al. Cancer statistics in China, 2015. CA Cancer J Clin. 2016;66(2):115-32.
3. Tamakoshi A, Nakamura K, Ukawa S, Okada E, Hirata M, Nagai A, et al. Characteristics and prognosis of Japanese colorectal cancer patients: The BioBank Japan Project. J Epidemiol. 2017;27(3s):S36-s42.
4. Bekaii-Saab TS, Ou FS, Ahn DH, Boland PM, Ciombor KK, Heying EN, et al. Regorafenib dose-optimisation in patients with refractory metastatic colorectal cancer (ReDOS): a randomised, multicentre, open-label, phase 2 study. Lancet Oncol. 2019;20(8):1070-82.
5. Mayer RJ, Van Cutsem E, Falcone A, Yoshino T, Garcia-Carbonero R, Mizunuma N, et al. Randomized trial of TAS-102 for refractory metastatic colorectal cancer. N Engl J Med. 2015;372(20):1909-19.
6. Sartore-Bianchi A, Trusolino L, Martino C, Bencardino K, Lonardi S, Bergamo F, et al. Dual-targeted therapy with trastuzumab and lapatinib in treatment-refractory, KRAS codon 12/13 wild-type, HER2-positive metastatic colorectal cancer (HERACLES): a proof-of-concept, multicentre, open-label, phase 2 trial. Lancet Oncol. 2016;17(6):738-46.
7. Mason JM, Arndt KM. Coiled coil domains: stability, specificity, and biological implications. Chembiochem. 2004;5(2):170-6.
8. Fan C, Dong L, Zhu N, Xiong Y, Zhang J, Wang L, et al. Isolation of siRNA target by biotinylated siRNA reveals that human CCDC12 promotes early erythroid differentiation. Leuk Res. 2012;36(6):779-83.
9. Schaafhausen A, Rost S, Oldenburg J, Muller CR. Identification of VKORC1 interaction partners by split-ubiquitin system and coimmunoprecipitation. Thromb Haemost. 2011;105(2):285-94.
10. Zhang X, Zheng Q, Wang C, Zhou H, Jiang G, Miao Y, et al. CCDC106 promotes non-small cell lung cancer cell proliferation. Oncotarget. 2017;8(16):26662-70.
11. Yin DT, Xu J, Lei M, Li H, Wang Y, Liu Z, et al. Characterization of the novel tumor-suppressor gene CCDC67 in papillary thyroid carcinoma. Oncotarget. 2016;7(5):5830-41.
12. Ke J, Lou J, Zhong R, Chen X, Li J, Liu C, et al. Identification of a Potential Regulatory Variant for Colorectal Cancer Risk Mapping to 3p21.31 in Chinese Population. Sci Rep. 2016;6:25194.
13. Li Q, Seo JH, Stranger B, McKenna A, Pe'er I, Laframboise T, et al. Integrative eQTL-based analyses reveal the biology of breast cancer risk loci. Cell. 2013;152(3):633-41.
14. Zhang J, Jiang K, Shen Z, Gao Z, Lv L, Ye Y, et al. Expression QTL-based analyses reveal the mechanisms underlying colorectal cancer predisposition. Tumour Biol. 2014;35(12):12607-11.
15. Manjelievskaia J, Brown D, McGlynn KA, Anderson W, Shriver CD, Zhu K. Chemotherapy Use and Survival Among Young and Middle-Aged Patients With Colon Cancer. JAMA Surg. 2017;152(5):452-9.
16. Lorenzon L, Biondi A, Carus T, Dziki A, Espin E, Figueiredo N, et al. Achieving high quality standards in laparoscopic colon resection for cancer: A Delphi consensus-based position paper. Eur J Surg Oncol. 2018;44(4):469-83.

17. Mirkin KA, Kulaylat AS, Hollenbeak CS, Messaris E. Robotic versus laparoscopic colectomy for stage I-III colon cancer: oncologic and long-term survival outcomes. Surg Endosc. 2018;32(6):2894-901.
18. Kuboki Y, Nishina T, Shinozaki E, Yamazaki K, Shitara K, Okamoto W, et al. TAS-102 plus bevacizumab for patients with metastatic colorectal cancer refractory to standard therapies (C-TASK FORCE): an investigator-initiated, open-label, single-arm, multicentre, phase 1/2 study. Lancet Oncol. 2017;18(9):1172-81.
19. Moriwaki T, Fukuoka S, Taniguchi H, Takashima A, Kumekawa Y, Kajiwara T, et al. Propensity Score Analysis of Regorafenib Versus Trifluridine/Tipiracil in Patients with Metastatic Colorectal Cancer Refractory to Standard Chemotherapy (REGOTAS): A Japanese Society for Cancer of the Colon and Rectum Multicenter Observational Study. Oncologist. 2018;23(1):7-15.
20. Hinds DA, Stokowski RP, Patil N, Konvicka K, Kershenobich D, Cox DR, et al. Matching strategies for genetic association studies in structured populations. Am J Hum Genet. 2004;74(2):317-25.
21. Lu X, Cao M, Han S, Yang Y, Zhou J. Colorectal cancer risk genes are functionally enriched in regulatory pathways. Sci Rep. 2016;6:25347.
22. Correll CC, Bartek J, Dundr M. The Nucleolus: A Multiphase Condensate Balancing Ribosome Synthesis and Translational Capacity in Health, Aging and Ribosomopathies. Cells. 2019;8(8).
23. Boulianne B, Feldhahn N. Transcribing malignancy: transcription-associated genomic instability in cancer. Oncogene. 2018;37(8):971-81.
24. Mladenov E, Magin S, Soni A, Iliakis G. DNA double-strand-break repair in higher eukaryotes and its role in genomic instability and cancer: Cell cycle and proliferation-dependent regulation. Semin Cancer Biol. 2016;37-38:51-64.
25. Moldogazieva NT, Lutsenko SV, Terentiev AA. Reactive Oxygen and Nitrogen Species-Induced Protein Modifications: Implication in Carcinogenesis and Anticancer Therapy. Cancer Res. 2018;78(21):6040-7.
26. Seelige R, Searles S, Bui JD. Mechanisms regulating immune surveillance of cellular stress in cancer. Cell Mol Life Sci. 2018;75(2):225-40.
27. Sengupta S, George RE. Super-Enhancer-Driven Transcriptional Dependencies in Cancer. Trends Cancer. 2017;3(4):269-81.
28. Studer S, Hansen DA, Pianowski ZL, Mittl PRE, Debon A, Guffy SL, et al. Evolution of a highly active and enantiospecific metalloenzyme from short peptides. Science. 2018;362(6420):1285-8.
29. Chasapis CT, Loutsidou AC, Spiliopoulou CA, Stefanidou ME. Zinc and human health: an update. Arch Toxicol. 2012;86(4):521-34.
30. Loh SN. The missing zinc: p53 misfolding and cancer. Metallomics. 2010;2(7):442-9.
31. Valencia AM, Kadoch C. Chromatin regulatory mechanisms and therapeutic opportunities in cancer. Nat Cell Biol. 2019;21(2):152-61.
32. Meldolesi J. Exosomes and Ectosomes in Intercellular Communication. Curr Biol. 2018;28(8):R435-r44.

33. Zhitnyak IY, Rubtsova SN, Litovka NI, Gloushankova NA. Early Events in Actin Cytoskeleton Dynamics and E-Cadherin-Mediated Cell-Cell Adhesion during Epithelial-Mesenchymal Transition. Cells. 2020;9(3).
34. Nieto MA, Huang RY, Jackson RA, Thiery JP. EMT: 2016. Cell. 2016;166(1):21-45.

Figures

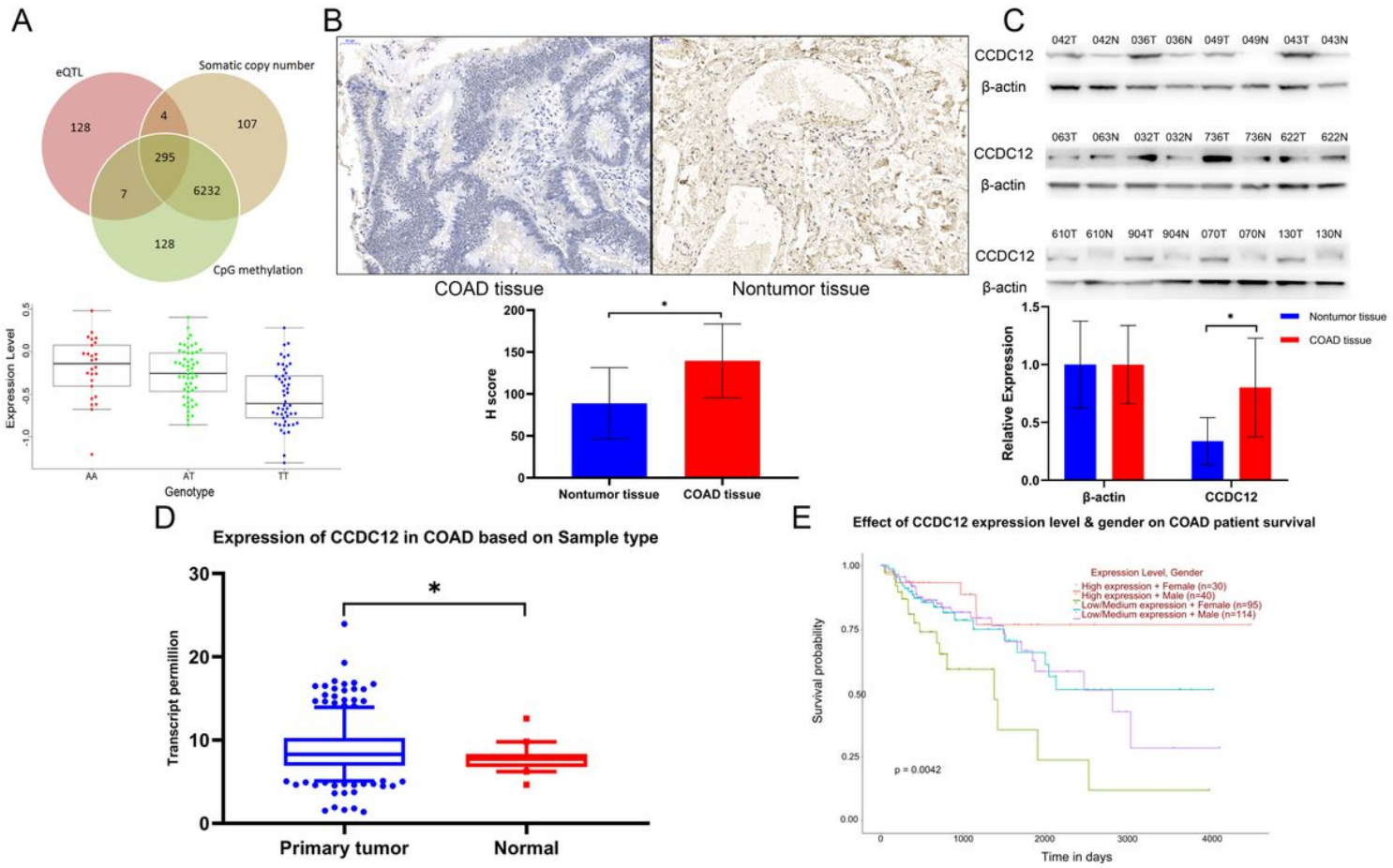

Figure 1

CCDC12 is an oncogene associated with colon cancer. (A) Cis-eQTL analysis determined a significant association between SNP rs8180040 and CCDC12 expression levels. (B) IHC staining demonstrating higher expression levels of CCDC12 in COAD tissues (40×). (C) Western blotting analysis demonstrating high expression levels of CCDC12 in COAD tissues. (D) The expression status of CCDC12 derived from COAD RNA-Seq data from TCGA database. (E) CCDC12 expression levels with different impact on the prognosis of COAD patients with gender classification.

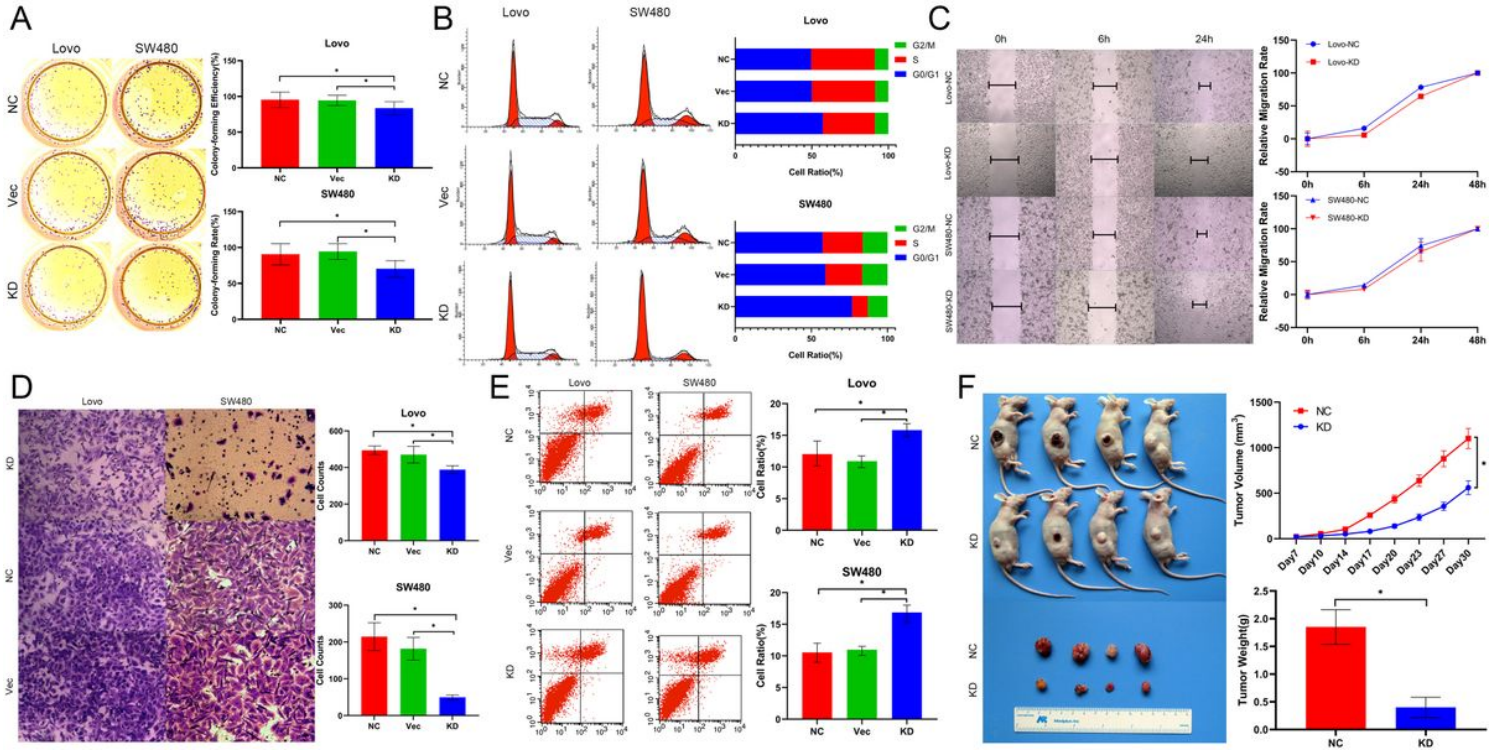

Figure 2

Knock-down of CCDC12 expression levels effectively inhibited the biological behavior of colon cancer cells. (A) Inhibition of colony formation in colon cancer cells (100×). (B) Changes in cell cycle phases after CCDC12 knockdown. (C) Decrease in distance of cell migration in CCDC12 knockdown cells (200×). (D) Reduced invasion of cells after CCDC12 knockdown (200×). (E) Apoptosis rates of cells. (F)Xenograft tumor volume and size were significantly reduced after knockdown of CCDC12.

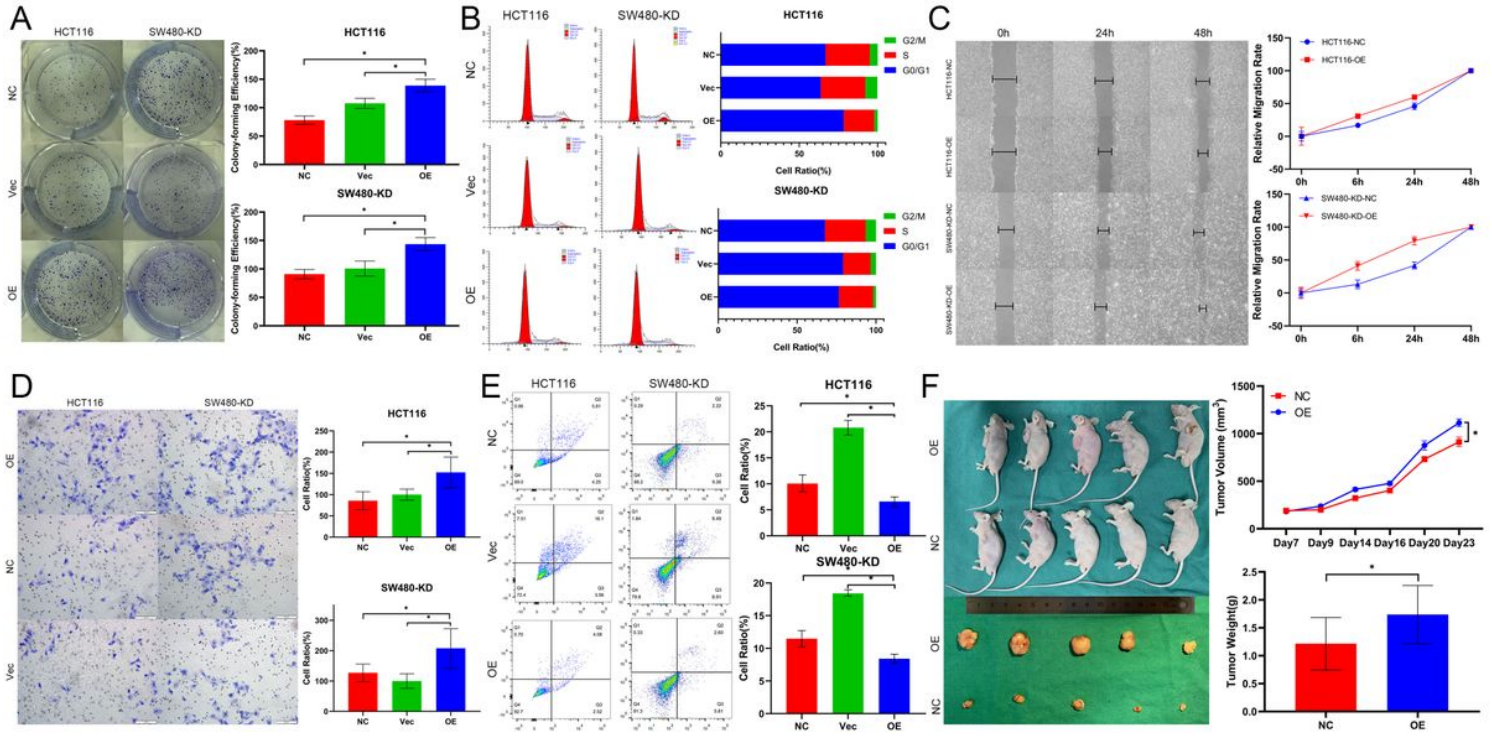

Figure 3

Overexpression CCDC12 enhances the biological behavior of tumor cells. (A) Increase in colony formation in colon cancer cells after CDCC12 overexpression (100×). (B) Changes in cell cycle phases after CCDC12 overexpression. (C) Increase in cell migration ability (200×). (D) Increase in invasion ability after CCDC12 overexpression (200×). (E) apoptosis assays. (F)Xenograft tumor volume and weight were significantly increased after CCDC12 overexpression.

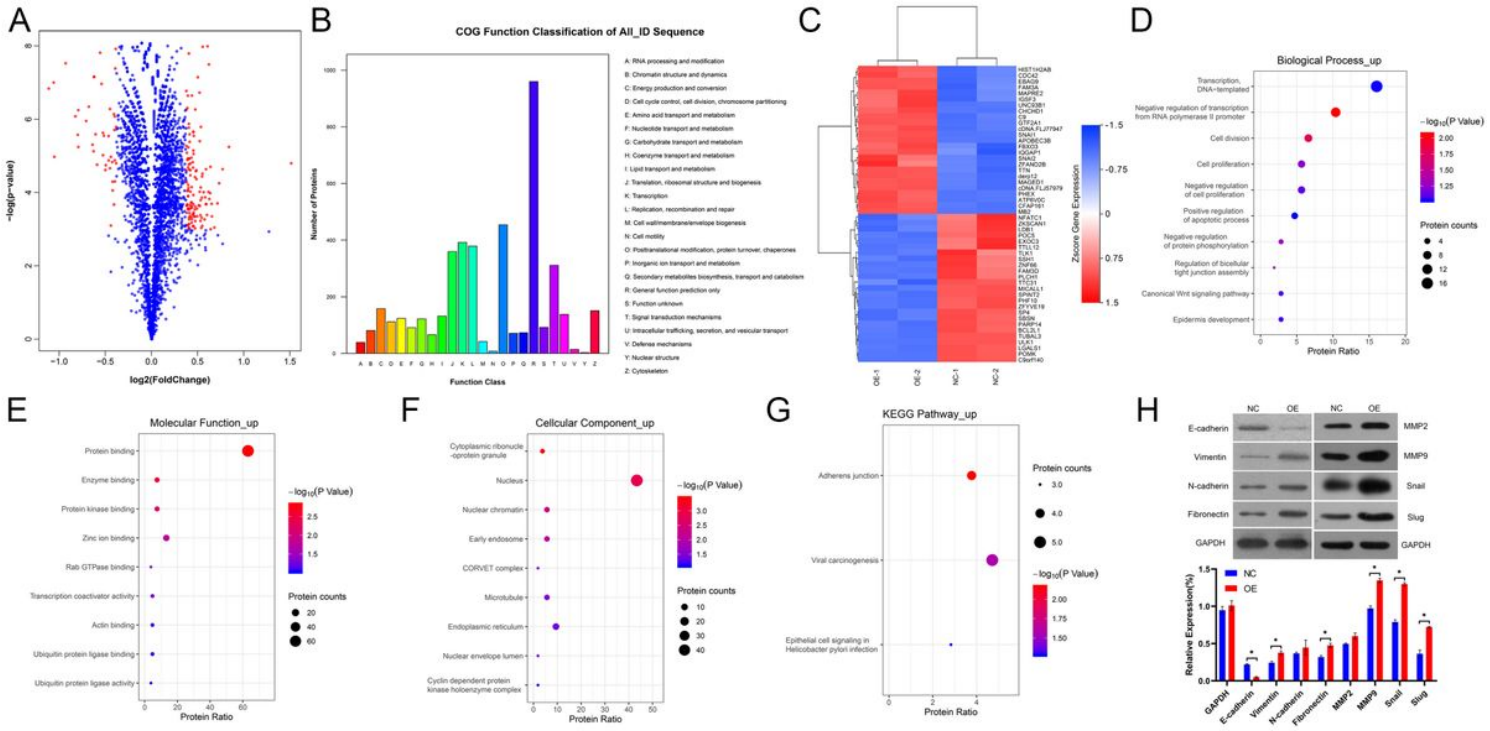

Figure 4

Bioinformatic analysis and validation of differentially expressed proteins (A) Volcano plot showing differentially expressed proteins. (B) COG annotation of protein function. (C) Cluster heatmap of the top 50 differentially expressed proteins. (D) GO annotation of the increased expressed proteins (BP). (E) GO annotation of the increased expressed proteins (MF). (F) GO annotation of the increased expressed proteins (CC). (G) KEGG signal pathway enrichment analysis. (H) Western blotting analysis to verify the association between CCDC12 and key EMT related proteins.

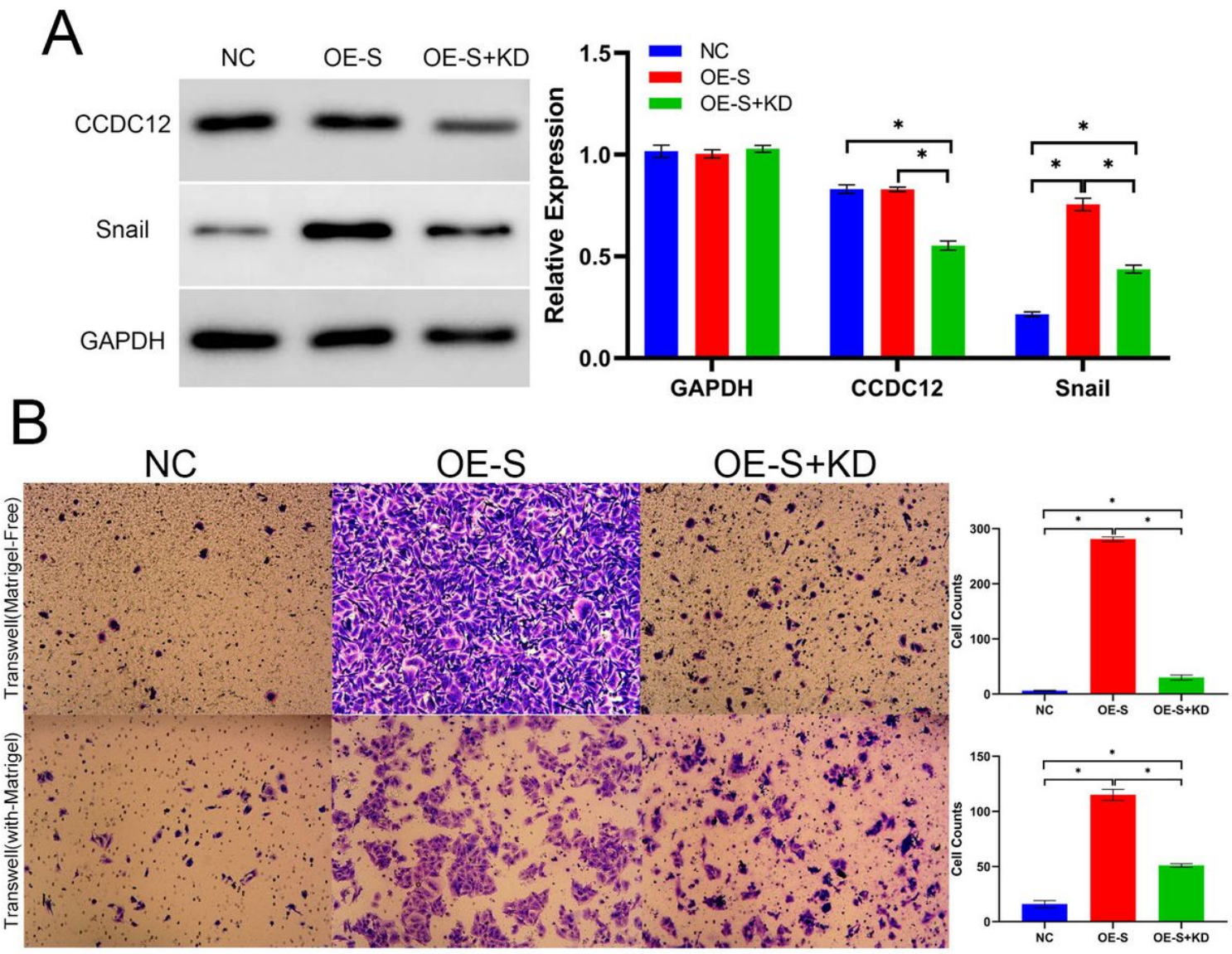

Figure 5

CCDC12 regulates colon cancer progression through Snail. (A) CCDC12 affects the expression levels of Snail. (B) CCDC12 regulates Snail to promote invasion and migration of colon cancer cells.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementarymaterial.docx
- Supplementarymaterial.docx