



ARTICLE

ADH1C inhibits progression of colorectal cancer through the ADH1C/PHGDH/PSAT1/serine metabolic pathway

Sha Li^{1,2}, Hong Yang^{1,2}, Wan Li^{1,2}, Jin-yi Liu^{1,2}, Li-wen Ren^{1,2}, Yi-hui Yang^{1,2}, Bin-bin Ge^{1,2}, Yi-zhi Zhang^{1,2}, Wei-qi Fu^{1,2}, Xiang-jin Zheng^{1,2}, Guan-hua Du^{1,2} and Jin-Hua Wang^{1,2}

Colorectal cancer (CRC) is the third most common cancer in men and the second most common cancer in women worldwide. CRC is the second leading cause of cancer-related deaths. Although some progress in the treatment of CRC has been achieved, the molecular mechanism of CRC is still unclear. In this study, alcohol dehydrogenase 1C (ADH1C) was first identified as a target gene closely associated with the development of CRC by the comprehensive application of transcriptomics, proteomics, metabolomics and in silico analysis. The ADH1C mRNA and protein expression in CRC cell lines and tumor tissues was lower than that in normal intestinal epithelial cell lines and healthy tissues. Overexpression of ADH1C inhibited the growth, migration, invasion and colony formation of CRC cell lines and prevented the growth of xenograft tumors in nude mice. The inhibitory effects of ADH1C on CRC cells in vitro were exerted by reducing the expression of PHGDH/PSAT1 and the serine level. This inhibition could be partially reversed by adding serine to the culture medium. These results showed that ADH1C is a potential drug target in CRC.

Keywords: colorectal cancer; ADH1C; metabolism; PHGDH; PSAT1; serine

Acta Pharmacologica Sinica (2022) 43:2709–2722; <https://doi.org/10.1038/s41401-022-00894-7>

INTRODUCTION

Cancer incidence and mortality are rapidly increasing worldwide. More than 1.2 million patients are diagnosed with colorectal cancer (CRC) every year, and more than 600,000 die from the disease [1]. CRC is the third most common cancer in men and the second most common cancer in women worldwide and the second leading cause of cancer-related deaths [2]. The incidence of CRC is associated with age, lifestyle, polyps in the colon, inflammatory bowel disease, family history and personal history and can also be considered a sign of socioeconomic development [3]. Abnormal metabolism is a hallmark of cancer, and more studies have shown that CRC is closely linked to abnormal metabolism [4].

Alcohol dehydrogenase 1C (ADH1C) is a member of the alcohol dehydrogenase family, which metabolizes ethanol, retinol and other aliphatic alcohols, hydroxysteroids, and lipid peroxidation products. Studies on the association of ADH1C with many types of cancers have provided significant but variable results, lacking consistency. ADH1C is associated with liver carcinoma [5, 6] and lung adenocarcinoma (LUAD) survival and is considered a promising biomarker for predicting the prognosis of LUAD [7]. In CRC, ADH1C is found to be involved in tumor immune cell infiltration and cetuximab resistance [8], while Ghosh et al. reported that ADH1B and ALDH2 but not ADH1C were associated with an increased risk of gastric cancer in West Bengal, India [9]. The ADH1C polymorphism may not be associated with the development of breast cancer in Caucasians [10] or with esophageal cancer in China [11]. The role of ADH1C in cancer is complex and controversial, and the mechanisms, especially in the case of CRC, remain unclear.

In the past decade, the molecular characterization of cancers via bioinformatics and multiomics analyses has helped elucidate cancer biology and identify new biomarkers and potential therapeutic targets. These methods provide researchers with highly efficient, powerful and reliable approaches to identify key genes related to the prognosis of various types of cancer [12, 13]. In this study, through the comprehensive application of transcriptomics, proteomics, metabolomics and in silico analyses, followed by verification with molecular biology analyses, we found that ADH1C was associated with CRC and explored its function in the progression of CRC. The results showed that the expression of ADH1C was downregulated in CRC cells and tissues and that overexpression of ADH1C inhibited the proliferation, migration, invasion, and colony formation of CRC cells. In addition, overexpression of ADH1C reduced the expression of two critical enzymes, phosphoglycerate dehydrogenase (PHGDH) and phosphoserine aminotransferase 1 (PSAT1), of the serine synthesis pathway (SSP). Serine is a nonessential amino acid that plays an important role in tumor metabolic reprogramming by participating in one-carbon metabolism, lipid metabolism, and protein biosynthesis. Knockdown of PHGDH or PSAT1 induced a phenotype of CRC cells caused by the overexpression of ADH1C, and this effect was reversed by exogenous serine. Furthermore, knockdown of PHGDH or PSAT1 reduced intracellular serine to the same extent as overexpression of ADH1C. Overexpression of ADH1C was also found to inhibit the growth of xenograft CRC tumors in vivo. In summary, ADH1C functions as a tumor suppressor gene, downregulating PHGDH and PSAT1 expression in the SSP, reducing intracellular serine and inhibiting CRC in vitro and in vivo.

¹The State Key Laboratory of Bioactive Substance and Function of Natural Medicines, Beijing 100050, China and ²Key Laboratory of Drug Target Research and Drug Screen, Institute of Materia Medica, Chinese Academy of Medical Science and Peking Union Medical College, Beijing 100050, China
Correspondence: Guan-hua Du (dugh@imm.ac.cn) or Jin-Hua Wang (wjh@imm.ac.cn)

Received: 30 October 2021 Accepted: 27 February 2022

Published online: 30 March 2022

MATERIALS AND METHODS

Identification and analysis of differentially expressed genes (DEGs) The datasets GSE33113, GSE21510, GSE9348 and GSE18105 were obtained from the platform at the GPL570 [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array and used to analyze the changes in mRNA expression levels between CRC and normal colorectal tissues acquired from GEO, one of the most widely used public databases for chip/gene profiles (<https://www.ncbi.nlm.nih.gov/geo/>). The GEO2R online tool [14] was used after replacing the matched probes with gene symbols of the original expression matrix, which was downloaded and used to find DEGs between CRC and normal tissues with an adjusted P value < 0.05 and $|\log_2FC| > 2$ among them (DEGs with $\log_2FC < 2$ are genes with significantly downregulated expression, while DEGs with $\log_2FC > 2$ are genes with significantly upregulated expression). Here, \log_2FC refers to $\log_2(\text{fold change})$. An online tool (<http://bioinformatics.psb.ugent.be/webtools/Venn/>) was used to draw the Venn diagram.

Weighted gene coexpression network analysis (WGCNA)

Following the developer's instructions [15], the WGCNA package (under R environment, version 3.6.3) was used to perform WGCNA on the expression matrix of GSE33113 and GSE9348. The scaleless adaptation index of different soft threshold parameters was precalculated using the `pickSoftThreshold` function. Converting the adjacency relationship to topological overlap indicates the network connectivity of a gene, which is defined as the sum of its adjacency relationship to all other genes for network generation. Based on TOM similarities and differences, the hierarchical clustering function was used to classify genes with similar expression profiles into modules, and the minimum size of the gene tree was 30. The dissimilarity of the intrinsic genes of the modules was calculated to select the cutoff point for merging modules. The network of intrinsic genes was visualized as a heatmap of 1000 randomly selected genes. The correlations between each module and clinical features were determined to identify the most relevant modules. Gene significance (GS) was defined as the \log_{10} transformation of the P value in linear regression between gene expression and clinical information ($GS = \log P$); module importance (MS) was defined as the average GS of all genes in the module. Generally, the module with the top MS ranking among all selected modules was regarded as a module related to clinical characteristics.

Enrichment analysis, protein-protein interaction (PPI), network construction, and hub gene identification

Gene Ontology (GO) enrichment analysis is a common method to define genes and their RNA or protein products to reveal the unique biological characteristics of high-throughput transcriptome or genomic data [16]. The Kyoto Encyclopedia of Genes and Genomes (KEGG) is a collection of databases dealing with genomes, diseases, biological pathways, drugs and chemical materials [17]. DAVID (database for annotation, visualization and integrated discovery) is an online bioinformatics tool designed to identify a large number of genes or protein functions [18] that was used to carry out GO analysis of biological processes (BPs), molecular functions (MFs), cellular components (CCs) and KEGG pathways ($P < 0.05$). These results were visualized as graphs using Circos [19].

The protein-protein interactions (PPIs) among the DEGs were generated using the STRING database in Cytoscape software (v.3.7.1) [20]. The hub genes, which are highly interconnected with the nodes in a group of DEGs or in a module of WGCNA, were identified by using the CytoHubba plug-in [21]. Using different algorithms at this step will result in different conclusions, so the top 20 hub genes and their scores were calculated by all eleven algorithms and comprehensively evaluated. The genes in a certain module with large numbers after WGCNA were filtered using their GS value and MM value ($|GS| > 0.2$ and $|MM| > 0.8$) to obtain the

specific number most likely to become hub genes from the algorithms mentioned above. The intersections between the hub genes from WGCNA and DEG analysis were assessed to obtain the final target genes. The transcriptional or post-transcriptional data on the differential expression of these target genes in CRC patient tumors and paracarcinoma tissues were obtained from the TCGA and GTEx RNA sequencing expression databases (<http://gepia.cancer-pku.cn/>), the OncoPrint database (<http://www.oncoPrint.org/>) and the Protein Atlas database (<https://www.proteinatlas.org/>).

Sources of cells and culture conditions

The human CRC cell lines HCT116, HCT8, HCT15, SW480, SW620, HT29, and T84 were purchased from the National Collection of Authenticated Cell Cultures (Shanghai, China). Cells were maintained in Dulbecco's Modified Eagle's Medium (DMEM, Gibco, NY, USA) supplemented with 10% fetal bovine serum [FBS, Gibco, streptomycin (100 $\mu\text{g}/\text{mL}$), and penicillin (100 U/mL)]. The cells were cultured at 37 °C in an incubator with a humidified atmosphere of 5% CO_2 .

Construction of ADH1C overexpression and knockdown cell lines

For construction of ADH1C-overexpressing CRC cell lines (ADH1C-OE), pCMV6-ADH1C-Myc-DDK (pADH1C) and pCMV6-Myc-DDK (pNC) were purchased from Origene (Beijing, China) and transfected into HCT116, SW620, HCT8 and HCT15 cells for 24 h, followed by seeding into a 10 cm dish with 1000 cells/dish. After culture for 24 h, G418 was added to a concentration of 1.2 mg/mL for screening. The culture medium was replaced with fresh medium containing 10% FBS and 1.2 mg/mL G418. Stable cell clones with high ADH1C expression were selected and cultured with DMEM containing 10% FBS and 1.2 mg/mL G418. siRNA duplexes were obtained from GenePharma (Shanghai, China) and transfected into HCT116, SW620, HCT8 and HCT15 cells using Lipofectamine™ 3000 (Thermo Fisher Scientific, Waltham, MA, USA) according to the manufacturer's instructions. The sequences of siRNA targeting PSAT1 were 5'-GCUGAAUGACAACACCUUUTT-3' and 5'-UAUAGUUGUCUGAACAGCTT-3'. The sequences of siRNA targeting the PHGDH gene were 5'-GCUGAAUGACAACACCUUUTT-3' and 5'-AAAGGUGUUGUCAUUCAGCTT-3'.

Cell proliferation

Cells (3000 per well) were seeded in 96-well plates and cultured at 37 °C for 24, 48 and 72 h, followed by incubation for 1 h with 10 μL of CCK-8 assay reagent (Beyotime, Shanghai, China). The absorbance at 450 nm was measured with a microplate reader (PerkinElmer, Envision 2104 multilabel reader, USA).

Migration and invasion assay

This experiment was performed using 6.5 mm Transwell® inserts of polycarbonate membranes with 8.0 μm pores (Corning, NY, USA). For the migration assay, 2×10^4 cells were seeded in the upper chamber and cultured at 37 °C for 6 h. The medium was then replaced with fresh, serum-free medium, and medium containing 20% FBS was added to the bottom chamber. After 24 h, the migrated cells were stained with 0.1% crystal violet (Solarbio, Beijing, China), and the total number of cells was counted under a fluorescence microscope (Nikon Eclipse Ti-U, Japan). For the invasion assay, the upper chamber was coated with Matrigel™ (Corning) diluted in serum-free medium at a ratio of 1:7. Aliquots of 1×10^5 cells were seeded in the upper chamber, cultured at 37 °C for 48 h, and then stained with 0.1% crystal violet.

Colony formation assay

Cells were suspended in medium containing 0.4% agar (Thermo Fisher Scientific) and plated onto a layer of 0.7% agar (5000 cells/well in 1.5 mL medium plus 2 mL bottom agar) in 6-well plates. After culture for two weeks, the colonies were stained with MTT (Sigma-Aldrich, St. Louis, USA) and counted.

Immunohistochemistry analysis of tissue microarrays (TMAs) CRC tissue microarrays (Cat No. OD-CT-DgLin01-001_1) were purchased from Outdo Biotech Co., Ltd. (Shanghai, China) containing tumor tissue and paired adjacent normal tissues from 93 CRC patients. After routine dewaxing and rehydration, antigen recovery was performed on the sections by microwaving in citric saline at 95 °C for 90 s, and endogenous peroxidase enzymes were neutralized with 3% hydrogen peroxide. After permeabilization with 0.1% Triton X-100 and blocking with 5% bovine serum albumin, the chip was incubated with 1:200 primary antibody against rabbit anti-ADH1C antibody (Thermo Fisher Scientific; Cat. No. PA5-76340) at 4 °C overnight. The microarray was equilibrated at room temperature for 30 min, washed with PBS and incubated with horseradish peroxidase-conjugated goat anti-rabbit IgG (Dako, Wuhan, China) for 60 min at room temperature. ADH1C expression was visualized with DAB substrate (Dako). The microarray was fully scanned by Panoramic Scan (3DHistech, Ltd., Hungary) and quantitatively analyzed using Image-Pro software (v.10.0.10, Media Cybernetics, Inc., MD, USA).

Quantitative real-time polymerase chain reaction (qRT-PCR)

After the indicated treatments, the cells were harvested in TRIzol (Ambion, Carlsbad, CA). After the cells were mixed with a 1/5 volume of chloroform (TongGuang, Beijing, China), the mixture was centrifuged at $12,000 \times g$ for 15 min, and the supernatants were transferred into new, clear centrifuge tubes. An equal volume of isopropanol (TongGuang) was added to each supernatant and gently mixed. After incubation at room temperature for 30 min, the mixture was centrifuged at $12,000 \times g$ for 15 min. The pellets were washed once with 75% ethanol (TongGuang) and dissolved in RNase-free water (Solarbio) at an appropriate volume. After RNA quantification, cDNA was synthesized using PrimeScript™ RT 1st master mix (TaKaRa, Japan) according to the manufacturer's instructions. Quantitative real-time RT-PCR (qRT-PCR) was performed using TB Green® premix ExTaq™II (Tli RNase H Plus) (TaKaRa). GAPDH served as an internal control. The primers used are listed in Supplementary Table 1.

Western immunoblotting (WB)

After the indicated treatments, the cells were harvested by centrifugation, and the pellets were resuspended in RIPA buffer (Applygen Technologies, Inc., Beijing, China) for protein extraction. The protein concentration was determined using a BCA assay kit from Applygen Technologies, Inc. Aliquots of 20–40 µg of protein were separated by 10% sodium dodecyl sulfate polyacrylamide gel electrophoresis and then transferred onto PVDF membranes (Merck Millipore, Ltd., Germany). The membranes were blocked with TBST containing 5% nonfat milk at room temperature for 2 h and incubated with GAPDH rabbit monoclonal antibody (Proteintech, Wuhan, China; Cat No. 6004-1-Ig), PSAT1 rabbit polyclonal antibody (Proteintech; Cat No. 10501-1-AP), PHGDH rabbit polyclonal antibody (Proteintech; Cat No. 14719-1-AP), or ADH1C rabbit polyclonal antibody (Thermo Fisher Scientific; Cat. No. PA5-76340) at 4 °C overnight. Subsequently, the membranes were washed three times with TBST and incubated with HRP-linked anti-mouse IgG antibody (Cell Signaling Technology, Danvers, USA; Cat. No. 7076S) or HRP-linked anti-rabbit IgG antibody (Cell Signaling Technology; Cat No. 7074S) at room temperature for 2 h. Finally, the membranes were washed three times with TBST and incubated with ECL reagents (CWBio, Beijing, China). The membranes were examined using a chemiluminescence photo-documentation system (Tanon, Beijing, China) and photographed.

Xenograft assay in nude mice

Animal studies were approved by the Committee on Animal Care & Welfare at the Institute of Materia Medica, Chinese Academy of Medical Science & Peking Union Medical College (No. 0004595). Approximately seven-week-old athymic nude mice (18–20 g) (Vital

River Laboratory Animal Technology Co., Ltd., Beijing, China) were housed in the animal facility in the Department of Animal Care Center at the Institute of Materia Medica, Chinese Academy of Medical Science & Peking Union Medical College at an ambient temperature of 24 °C with unlimited access to food and water. All experimental procedures were carried out in accordance with institutional guidelines for the care and use of laboratory animals at the Institute of Materia Medica, Chinese Academy of Medical Science & Peking Union Medical College and the National Institutes of Health Guide for Care and Use of Laboratory Animals (publication No. 85-23, revised 1985). Mice were randomly distributed at six per group, and an aliquot of HCT116 or HCT116_ADH1C-OE cells was subcutaneously injected into the right flank of each mouse. Tumor volume (mm^3) was measured with a Vernier caliper and calculated using the formula $(L \times W^2)/2$, where L and W are the length and width of the tumor, respectively.

Transcriptomics

HCT116 and HCT116_ADH1C-OE cells were harvested in TRIzol for RNA isolation and sequencing. Initial isolates were assessed for quality by FastQC (v.0.11.9, open source) and filtered to remove low-quality calls using default parameters and specifying a minimum length of 50. Processed reads were then aligned to the *Homo sapiens* genome assembly with Cuffmerge (Tuxedo Suite pipeline). The mRNA levels were evaluated as fragments per kB per million fragments (FPKM) using Cuffquant and Cuffnorm (Tuxedo Suite pipeline). The sample correlation analysis was performed using the Pearson coefficient. Cuffdiff (Tuxedo Suite pipeline) was used to analyze the differential expression, and the default screening standard for differential expression was $|\log_2\text{FC}| \geq 1$ and P value ≤ 0.05 . Genes with $\log_2\text{FC} \geq 1$ represented those with upregulated expression, while $\log_2\text{FC} \leq -1$ represented those with downregulated expression. The common differentially expressed genes from HCT116 and HCT116_ADH1C-OE cells were collected for subsequent analysis. Gene enrichment analyses and PPI analyses were performed as described above.

Proteomics—extraction and analysis by LC-MS

Proteins of HCT116 and HCT116_ADH1C-OE cells were extracted with 8 M urea, reduced with 10 mmol/L dithiothreitol, and alkylated with 55 mmol/L iodoacetamide. Proteins were adjusted to equivalent concentrations with 20 mmol/L Tris-HCl buffer in 30 kD ultracentrifugation tubes and digested with 0.25% trypsin (Solarbio) overnight at 37 °C in ultracentrifugation tubes. The tryptic peptides were desalted and concentrated using reversed-phase C18 STAGE tips. The elution products were dried in a vacuum centrifuge to remove solvent. Peptides were dissolved in 0.1% formic acid and separated using an EASY-n LC1000 system. The column oven was set to 60 °C. The peptides were delivered to a trap column (75 µm × 2 cm, C18, 5 µm, Thermo Scientific) and then separated with a capillary LC column (75 µm × 100 mm, C18, 3 µm, Kyoto Monotech). The elution gradient was 6%–28% for 48 min and 28%–95% for 4 min (buffer B = 0.1% formic acid, 100% ACN; flow rate, 0.6 µL/min). An Orbitrap Fusion Lumos mass spectrometer was used to analyze the eluted peptides from LC. The data were obtained by data-independent acquisition under the high-sensitivity mode using the following parameters: positive mode was set; one cycle contained one full scan and 19 fragment scans; the full scan range was from 350 to 1300 m/z and screened at 120,000 resolution; fragment spectra were collected at 3,000 resolution and segmented to 19 MS/MS scans. The maximum injection time was 50 milliseconds. Raw DIA data were analyzed by Spectronaut (v.14.3, Biognosys, Switzerland) with default settings. The spectral libraries of DIA were generated using total raw data files with a Q value cutoff of 0.01 and a minimum of six fragment ions. The raw files were searched against the human database downloaded from the reviewed SwissProt database. Decoy items were generated by inverse mode.

The samples were subjected to quantitative evaluation based on the MS2 area. Cross runs were normalized according to the global abundance area.

Metabonomics

HCT116 and HCT116_ADH1C-OE cells were harvested and mixed with 10 prechilled zirconium oxide beads and 20 μL of deionized water. The samples were homogenized for 3 min, and 150 μL of methanol containing internal standard was added to extract the metabolites. The samples were homogenized for another 3 min and then centrifuged at $18,000 \times g$ for 20 min. The supernatants were transferred to a 96-well plate, and the following procedures were performed on an Eppendorf epMotion workstation (Eppendorf, Inc., Hamburg, Germany). Freshly prepared derivatization reagents (20 μL) were added to each well. The plate was sealed, and the derivatization was carried out at 30 $^{\circ}\text{C}$ for 60 min. After derivatization, the samples were evaporated for 2 h, and 330 μL of ice-cold 50% methanol solution was added to reconstitute the sample. The plate was held at -20°C for 20 min and then centrifuged at $4000 \times g$ for 30 min at 4 $^{\circ}\text{C}$. Aliquots of supernatants (135 μL) were transferred to a new 96-well plate with 10 μL of internal standards in each well. Serial dilutions of derivatized stock standards were added to the left wells. An ultra-performance liquid chromatograph coupled to a tandem mass spectrometry (UPLC-MS/MS) system (ACQUITY UPLC-Xevo TQ-S, Waters Corp., Milford, USA) was used to quantitate the targeted metabolites. The raw data files generated by UPLC-MS/MS were processed using MassLynx software (v4.1, Waters, Milford, USA) to perform peak integration, calibration, and quantitation for each metabolite. The self-developed platform lmap (v1.0, MetaboProfile, Shanghai, China) was used for statistical analyses.

Statistical analysis

All experiments were repeated three times, and the data are expressed as the mean \pm SD. Statistical significance between different groups was analyzed by one-way ANOVA using GraphPad software (v.7.00, GraphPad Software, Inc., San Diego, California, USA). The D'Agostino-Pearson test was applied to determine if the values followed a Gaussian distribution. $P < 0.05$ was considered statistically significant.

RESULTS

Identification of target genes in CRC by bioinformatics

To identify genes associated with the development of CRC, four datasets (GSE33113, GSE21510, GSE9348, GSE18105) from the GEO database, which contained mRNA expression data of 358 CRC tissues and 79 normal colorectal tissues were selected and analyzed. Using the online GEO2R tool, 1570, 2860, 451 and 3450 DEGs were obtained from GSE33113, GSE9348, GSE21510 and GSE18105, respectively. Among these DEGs, 815, 2187, 236 and 3079 genes were upregulated and 755, 673, 215 and 371 genes were downregulated ($P < 0.05$, $|\log_2\text{FC}| > 2$). The Venn diagram tool was used to show the intersection of the DEGs in the four datasets, and 96 upregulated DEGs and 110 downregulated DEGs were obtained (Fig. 1a). More details of these DEGs are shown in Supplementary Fig. 1.

WGCNA was conducted on the four GEO datasets to construct a coexpression network. Based on the Euclidean distance calculated by the \log_{10} -converted RNA-seq score counts, the samples were hierarchically clustered. In addition, the basic patient information was added under the generated tree (Supplementary Fig. 2a). The sample cluster results showed no obvious outliers. The proper soft threshold verification converged toward a scale-free topology, identifying values of 5 and 3 of GSE33113 and GSE9348, respectively (Supplementary Fig. 2b). The other two datasets, GSE18105 and GSE21510, were discarded because there was no suitable soft threshold or the soft threshold was too large, which

might have seriously distorted the analytical results. As shown in Fig. 1b, the coexpression network was constructed by the one-step method using the dynamic tree cut algorithm to obtain a group of modules in which the expression of member genes changed in a similar pattern. In GSE33113 (left panel) and GSE9348 (right panel), 15 and 20 modules were obtained, and the specific module size is shown at the bottom of Fig. 1b. Next, the interaction of the modules was analyzed, and a network heatmap between the modules (Supplementary Fig. 2c) and a network heatmap of 1000 randomly selected genes (Supplementary Fig. 2d) were plotted. These results showed that each module has a high degree of independence, and the gene expression in each module is relatively independent. For each module, the coexpression of genes was summarized by the intrinsic gene (the expression of the first component of the gene belonging to the module), and the relevance between each intrinsic gene and clinical characteristics was determined (tumorigenesis, recurrence, sex, age, etc.). There were four key modules most relevant to tumorigenesis in GSE33113 and GSE9348. As shown in Fig. 1c, the downregulated yellow module and the upregulated turquoise module were closely related to tumorigenesis in GSE33113, while the downregulated turquoise module and the upregulated brown module were most related to tumorigenesis in GSE9348. The gene significance (GS) across modules of the two datasets is shown in Supplementary Fig. 2e. Scatterplots (Supplementary Fig. 2f, g) among these four specific modules indicated that there was a highly significant correlation between GS and module membership in each module. These results allowed us to select them as key modules for further analysis.

GO enrichment analysis and KEGG pathway analysis were next performed on the gene clusters obtained by DEG identification and WGCNA, respectively. The biological processes (BPs) in which these gene clusters were involved and the specific genes involved in these BPs are shown in Fig. 1d. The results of the molecular function (MF), cell component (CC), and KEGG pathway analyses are shown in Supplementary Fig. 3. The PPI networks were constructed and evaluated comprehensively using the scores provided by all available algorithms in CytoHubba (Fig. 1e, f). As a result, 39 genes with upregulated expression and 38 genes with downregulated expression were identified from DEG analysis, 49 genes with upregulated expression and 61 genes with downregulated expression were identified from GSE33113, and 36 genes with upregulated expression and 60 genes with downregulated expression were identified from GSE9348. The details are listed in Supplementary Table 2.

The intersection between hub genes with upregulated or downregulated expression obtained in the DEG analysis and GSE33113 and GSE9348 was examined. The genes with upregulated expression in GSE33113 (turquoise module) and GSE9348 (brown module) had three common hub genes, namely, CEP55, ANLN and CCNB1, which were not included in the hub genes of the DEGs (Fig. 1g, left panel). The five common hub genes, MS4A12, GUCA2A, GUCA2B, ADH1C, and CLCA4, which appeared in all three gene clusters at the same time, were considered valuable target genes for further research (Fig. 1g, right panel).

To further validate the five target genes in CRCs, many famous cancers datasets including TCGA, Oncomine and Protein Atlas were used. According to the TCGA database, the five target genes had significantly downregulated expression in CRC (Fig. 2a), and the results from Oncomine supported this conclusion (Fig. 2b). In addition, the IHC staining results from the Protein Atlas database confirmed that changes in the protein expression of the hub genes were related to CRC tumorigenesis (Fig. 2c). Interestingly, the expression of ADH1C, CLCA4 and MS4A12 was correlated with the overall survival of CRC patients in TCGA databases (Fig. 2d). Thus, these five target genes, especially ADH1C, CLCA4 and MS4A12, play an important role in CRC tumorigenesis.

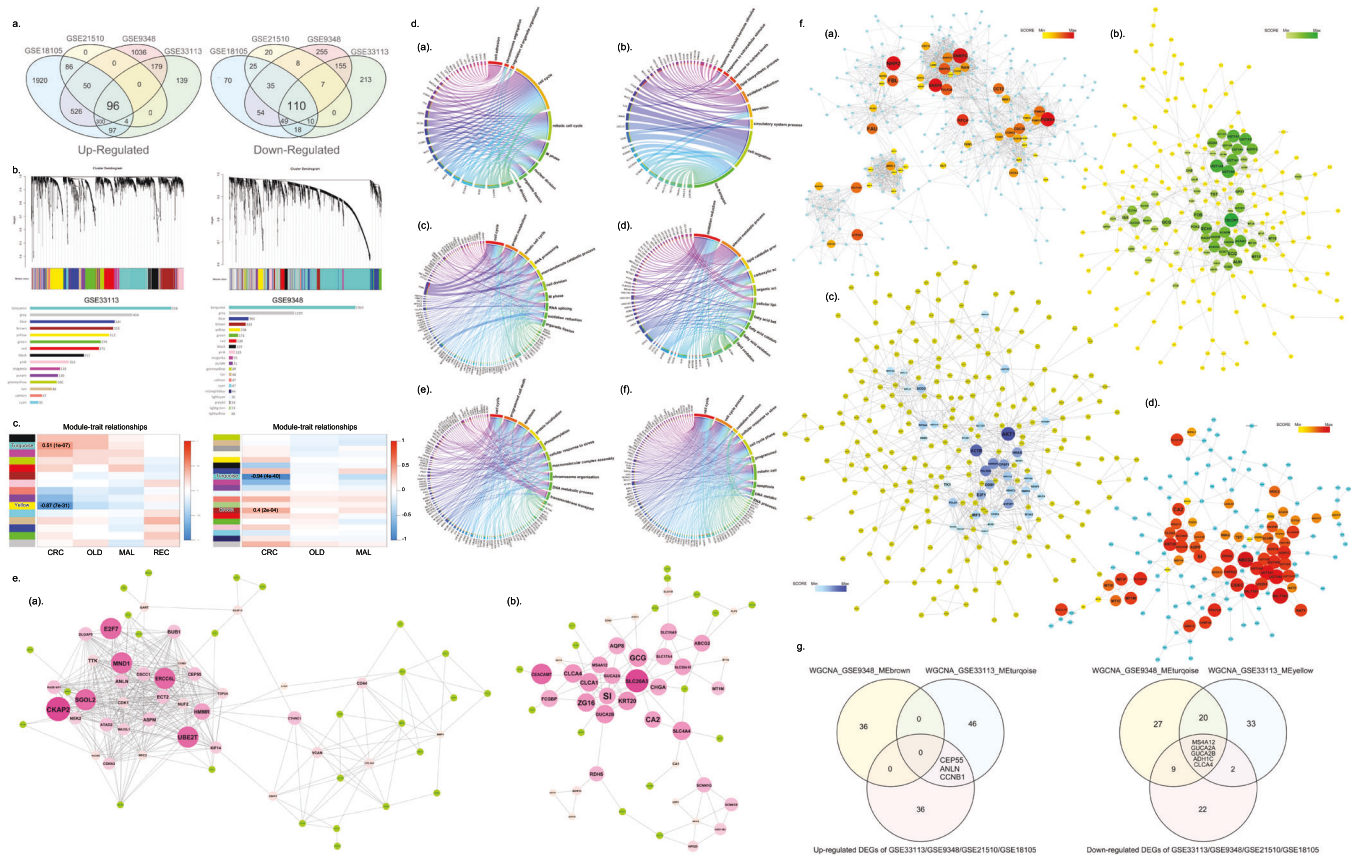


Fig. 1 Identification of the hub genes in CRC by DEGs analysis and WGCNA. **a** The DEGs of the four datasets and their intersection. **b** GSE3113 and GSE9348 modules have identified hierarchical cluster dendrograms of co-expressed genes. Each color line represents a color coding module, which contains a set of highly connected genes. Genes that cannot be accepted by any other module are indicated by gray modules. A total of 15 and 20 modules were identified in GSE3113 and GSE9348, respectively. The lower panel lists the number of genes contained in each module. **c** Heat map of the correlation between the intrinsic genes of each module and clinical features in GSE3113 and GSE9348. Each row corresponds to the module characteristic gene, and each column corresponds to the clinical characteristic. **d** Biological process (BP) of upregulated genes (a) and downregulated genes (b) in DEGs analysis. BP of GSE3113 turquoise module (c), GSE3113 yellow module (d), GSE9348 brown module (e) and GSE9348 turquoise module (f) in WGCNA. **e** PPI network and hub gene of upregulated (a) and downregulated (b) DEGs. Circles represent the nodes in the PPI network, gray straight lines represent the connection between the nodes. Pink circles represent the calculated hub genes, the circle becomes larger and the color becomes darker as the comprehensive scores increase. **f** PPI networks and hub genes of the turquoise module (a) and the yellow module (b) in GSE3113. PPI networks and hub genes of the brown module (c) and the turquoise module (d) in GSE9348. **g** Left panel: the intersection of GSE3113, GSE9348 and upregulated expression of DEGs. Right panel: the intersection of GSE3113, GSE9348 and downregulated expression of DEGs.

ADH1C mRNA and protein levels are downregulated in CRC tissues and cells

To focus on the role and mechanism of ADH1C in CRC, we first analyzed the expression of ADH1C *in silico*. There was much lower ADH1C mRNA expression in the CRC tissues than in the normal tissues (Fig. 3a). Then, ADH1C mRNA and protein expression levels were assessed by qRT-PCR and WB in the CRC cell lines HCT116, HCT8, HCT15, HT29, SW620, SW480 and T84. The results showed that ADH1C expression level was lower in the CRC cell lines than in the normal intestinal epithelium cell line NCM 460 (Fig. 3b, c). To further confirm the changes in ADH1C protein expression in CRC, IHC assays were carried out on CRC tissue microarrays. The ADH1C protein expression in the CRC tissues was lower than that in the paired healthy paracarcinoma tissues (Fig. 3d). The analyzed areas of each pair of tissues are shown in Supplementary Fig. 4. Collectively, these results showed that the expression of ADH1C mRNA and protein was lower in the CRC tissues than in the normal tissues.

ADH1C functions as a CRC suppressor *in vitro*

To explore the functional role of ADH1C in CRC, cell proliferation, migration, invasion and colony formation assays were performed

in cells with stable high ADH1C expression and control CRC cells. Overexpression of ADH1C inhibited the growth (Fig. 4a and Supplementary Fig. 5a) and colony formation of HCT116, SW620, HCT8 and HCT15 cells (Fig. 4b, Supplementary Fig. 5b). In addition, overexpression of ADH1C reduced the migration (Fig. 4c, d) and invasion (Fig. 4c) of HCT116 and SW620 cells. Altogether, ADH1C plays a key inhibitory role in the development of CRC.

Multiomics reveals the mechanism by which ADH1C functions as a tumor suppressor in CRC

To further elucidate the mechanism of ADH1C inhibition of CRC, multi-omics assays (transcriptomics, proteomics, metabolomics) were carried out. By comparing transcriptomics and proteomics data of HCT116 cells and HCT116 cells with high ADH1C expression, we demonstrated that the transcription of 183 genes was significantly changed ($|\log_2FC| > 1.5$), and the translation of 1569 proteins was significantly altered ($|\log_2FC| > 1.5$). More details of the results of the enrichment analysis are shown in Supplementary Fig. 6. There were 15 members (ADH1C, ASNS, HSPA1A, PSAT1, PHGDH, U2AF1L5, ERF1, HSPA8, HSPA6, TUBA4A, PYGB, PCK2, NIT1, HSPH1, and ARHGEF2) in the intersection, and four of them (ADH1C, ASNS, HSPA1A, and PSAT1) were proven to be

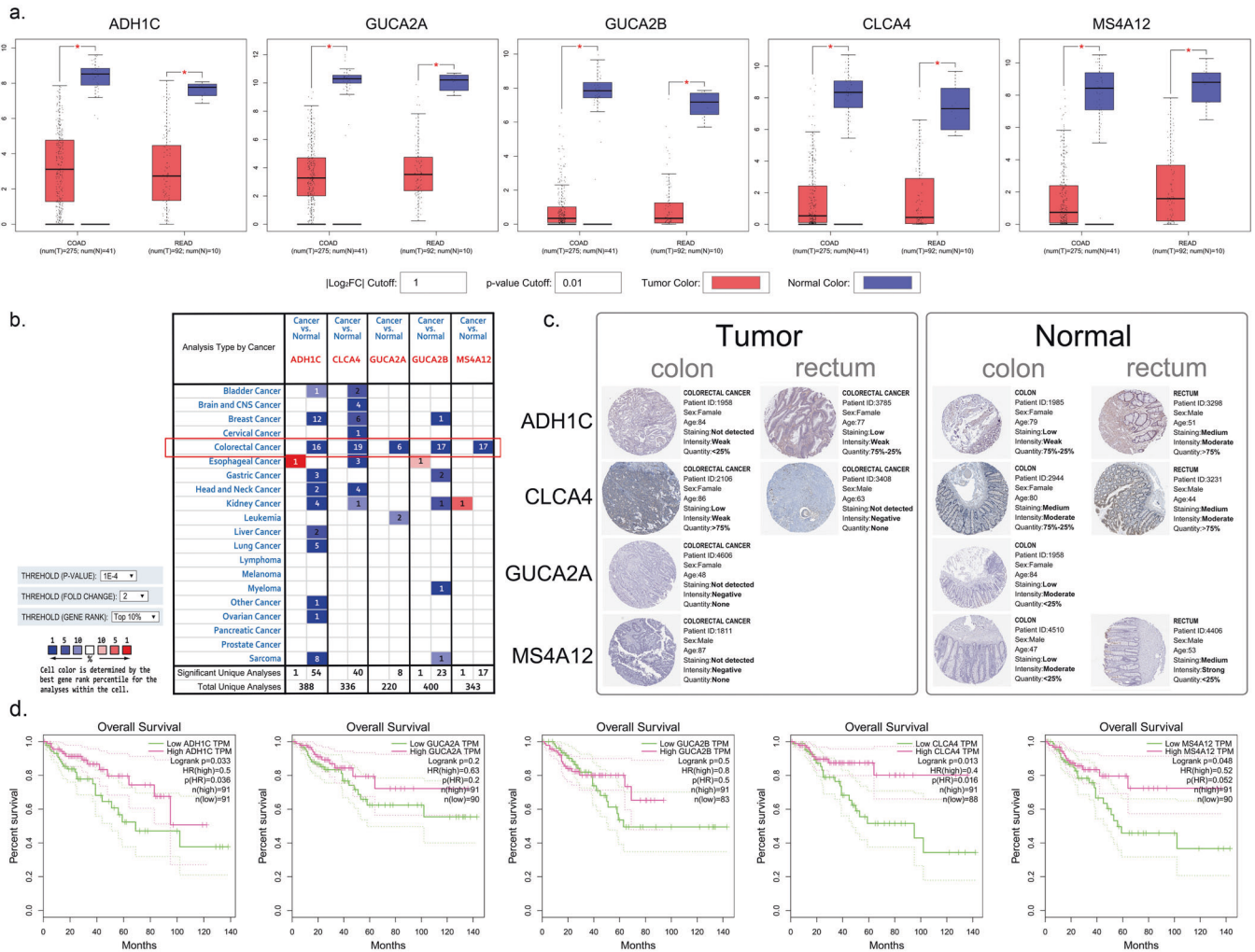


Fig. 2 Validation of candidate genes by bioinformatics methods. The mRNA expression level of the candidate genes in **a** TCGA and GTEx databases and **b** OncoPrint database. **c** The protein expression level of the candidate genes in Protein Atlas database (GUCA2B is not provided in the database). There is lower protein expression of the candidate genes in colorectal tumors than that in normal colon and rectal tissues. **d** Among the five candidate genes, the expression levels of ADH1C, CLCA4, and MS4A12 are associated with the overall survival of CRC according to TCGA and GTEx databases

downregulated and related to overall survival in CRC patients (Fig. 5a; Supplementary Fig. 7a, b). These results were verified by qRT-PCR (Fig. 5b). PHGDH is an enzyme that catalyzes the synthesis of serine, and high PHGDH expression was associated with poor outcome in CRC patients [22]. PSAT1 is another enzyme involved in serine biosynthesis that promotes the growth of cells and increases the chemoresistance of colon cancer cells [23]. Asparagine synthetase (ASNS), an enzyme that synthesizes asparagine from aspartate [24], catalyzes the ATP-dependent conversion of aspartate to asparagine (Fig. 5c). Because these potential research targets are related to metabolic processes, we next assessed the effect of ADH1C on a wide range of metabolites, including amino acids, by carrying out a metabolomics assay. Overexpression of ADH1C led to a significant decrease in intracellular asparagine and serine, which was consistent with the results of transcriptomics and proteomics (Fig. 5d). Taken together, these results suggested that ADH1C might play an inhibitory role in the development of CRCs by inhibiting the expression of the serine synthase PHGDH and PSAT1.

ADH1C suppresses CRC by inhibiting the phosphatidylserine synthases PHGDH and PSAT1

To confirm the hypothesis that ADH1C plays an inhibitory role in the development of CRC by inhibiting the expression of the serine

synthase PHGDH and PSAT1, expression of PHGDH and PSAT1 was measured in CRC cell lines with high ADH1C expression by WB. Overexpression of ADH1C significantly reduced the expression of PHGDH and PSAT1 in HCT116, HCT8, HCT15 and SW620 cells (Fig. 6a and Supplementary Fig. 8a). When the expression of PHGDH and PSAT1 was knocked down by siRNA (Fig. 6b and Supplementary Fig. 8a) and the functional role of ADH1C in CRC cells was investigated, we found that knockdown of PHGDH and PSAT1 expression inhibited cell proliferation (Fig. 6c, Supplementary Fig. 8b) and reduced migration (Fig. 6d, e, Supplementary Fig. 8c, d), invasion (Fig. 6d, Supplementary Fig. 8c), and colony formation of CRC cells (Fig. 6f, Supplementary Fig. 8e). These effects are consistent with the effects caused by overexpression of ADH1C. In addition, it was found that both ADH1C overexpression and PHGDH or PSAT1 knockdown significantly reduced the intracellular serine content in CRC cells and that there was no significant difference between them. It was shown in Fig. 6g and Supplementary Fig. 8f that serine biosynthesis was inhibited both by overexpression of ADH1C and by knockdown of PHGDH or PSAT1 expression in CRC cells. Addition of 12 mM serine to the culture medium reversed the effects of overexpression of ADH1C and knockdown of PHGDH or PSAT1 expression in CRC cells, which further proved that the effect of ADH1C on CRC cells was achieved

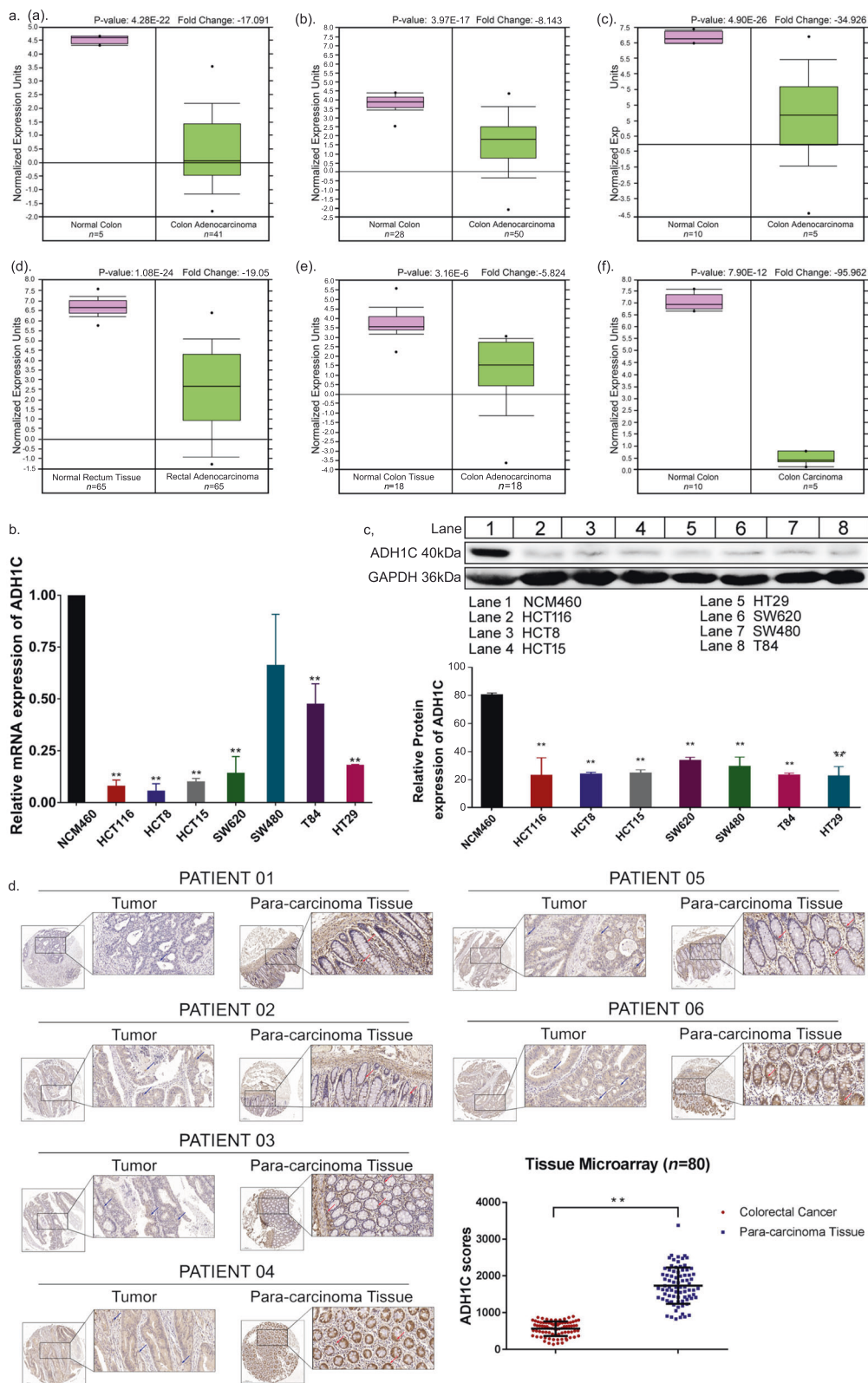


Fig. 3 ADH1C expression is negatively related with CRC. **a** There was lower ADH1C expression in colorectal tumors than that in para-carcinoma tissue. Data and statistics were obtained from www.oncomine.org. (a) Kaiser et al. (2007); (b) Ki et al. (2007); (c) Hong et al. (2010); (d) Gaedcke et al. (2010); (e) Nortterman et al. (2010); (f) Skrzypczak et al. (2010). Compared with normal colorectal cell line NCM460, ADH1C expression is lower in CRC cell lines at both **b** transcriptional level and **c** post-transcriptional level. **d** There was lower expression of ADH1C protein in tumors than that in para-carcinoma tissue of CRC tissue microarray. * $P < 0.05$, ** $P < 0.01$.

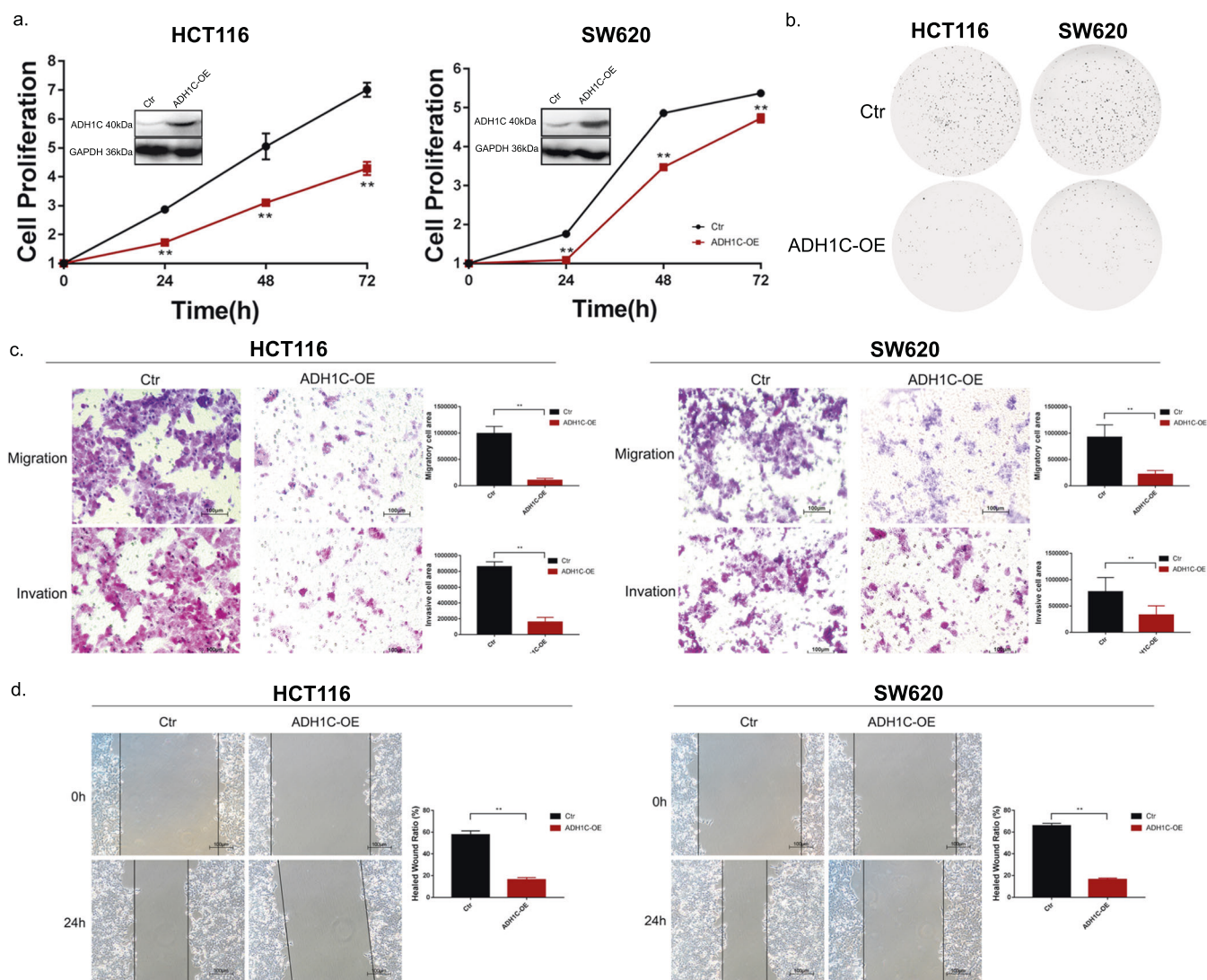


Fig. 4 ADH1C functions as a CRC suppressor in vitro. **a** Overexpression of ADH1C inhibited proliferation of CRC cell lines. **b** Overexpression of ADH1C reduced colony formation of CRC cell lines. **c** Overexpression of ADH1C inhibited migration and invasion of CRC cell lines. **d** Overexpression of ADH1C reduced migration (wound healing method) of CRC cell lines. ** $P < 0.01$ compared with control.

partially by inhibiting serine synthesis to reduce serine biosynthesis (Fig. 6h, Supplementary Fig. 8g). Our study also showed that adding 12 mM serine to ordinary DMEM was sufficient to significantly increase the proliferation of HCT116 cells (Supplementary Fig. 7c).

ADH1C functions as a CRC suppressor in vivo

To determine whether overexpression of ADH1C could inhibit the growth of xenografted CRC tumors in vivo, a xenograft CRC tumor model was established in immunodeficient mice by injecting 5.5×10^6 HCT116 cells with high ADH1C expression. The tumor volumes of the mice were determined every three days, and once tumors were palpable at $\sim 100 \text{ mm}^3$ in volume, the mice were euthanized. As shown in Fig. 7a–c, overexpression of ADH1C resulted in a significant decrease in tumor size and a lower average ratio of tumor weight and body weight. A statistically significant reduction in tumor volume occurred 4 days after the mice received CRC cells (Fig. 7d). Overexpression of ADH1C had no side effects on the major organs of the test mice (Fig. 7e). Overall, these results demonstrate that ADH1C inhibits the growth of xenograft tumors in vivo.

DISCUSSION

The cornerstones of CRC therapy are surgery, neoadjuvant radiotherapy (for patients with rectal cancer), and adjuvant chemotherapy (for patients with stage III/IV and high-risk stage II colon cancer) [1]. Although several targeted drugs have doubled the overall survival of patients with advanced disease (up to three years), CRC is still associated with a poor prognosis and a very low long-term survival rate [25]. The molecular pathogenesis of CRC is heterogeneous. The adenoma–carcinoma sequence, inherited forms, mismatch repair deficiency and microsatellite instability contribute to CRC [1]. In addition, the gut microbiota is considered important in the initiation and development of CRC [25]. However, as a disease closely related to metabolism, the exact molecular mechanism of CRC, especially the mechanism related to metabolism, remains unclear.

In recent years, DEG analysis based on a large amount of gene chip data in public databases has been one of the most widely used bioinformatics methods in tumor studies [26, 27]. However, the process of DEG analysis in which a group of genes that conform to certain standards (such as $P < 0.05$, $|\log_2\text{FC}| > 2$, etc.) are filtered out before calculating their connectivity and coexpression will reveal an



Fig. 5 Multi-omics reveals the mechanism by which ADH1C functions as a tumor-suppressor in CRC. **a** There were 183 genes and 1569 proteins which were significantly changed in HCT116 cells with overexpressing ADH1C ($|\log_2FC| > 1.5$). There are 15 members in their intersection and four of them, ADH1C, ASNS, HSPA1A, PSAT1, were proved to be associated with overall survival of CRC patients. **b** Verification of partial genes with downregulated expression in the transcriptome results. **c** The biosynthesis of serine and glycine from glucose, PHGDH and PSAT1 in the serine synthesis pathway (left); the biosynthesis of asparagine and the role of ASNS. **d** Metabolomics also showed that overexpression of ADH1C significantly reduced intracellular asparagine and serine in CRC cells with overexpression of ADH1C. PHGDH phosphoglycerate dehydrogenase; PSAT1 phosphoserine aminotransferase; ASNS asparagine synthetase

association of some genes whose expression differences were missed because they were not obvious but were also highly involved in the coexpression network. WGCNA is another tool [15] that can analyze tumor-related gene expression data from a different perspective. WGCNA calculates the expression relationship between genes, identifies gene modules with similar expression patterns, analyzes the relationship between gene sets and sample phenotypes, draws regulatory networks between genes in gene sets and identifies key regulatory genes. The advantages of WGCNA, such as making full use of information, converting the association of thousands of genes and phenotypes into the association of several gene sets and phenotypes, and eliminating the problem of multiple

hypothesis test correction, make it possible to successfully analyze the relationship between gene expression data and clinical features in various types of cancer, such as lung cancer [28], gastric cancer [29], glioma [30], pancreatic cancer [31] and bladder cancer [32]. By intersecting the hub genes with high connectivity of interacting proteins identified by constructing a PPI network for genes with certain differential expression patterns after performing DEG analysis and WGCNA, we identified CRC-related target genes that simultaneously possess several important characteristics of coexpression, differential expression, and protein interaction. ADH1C, CLCA4 and MS4A12 were selected by combining DEG analysis with WGCNA.

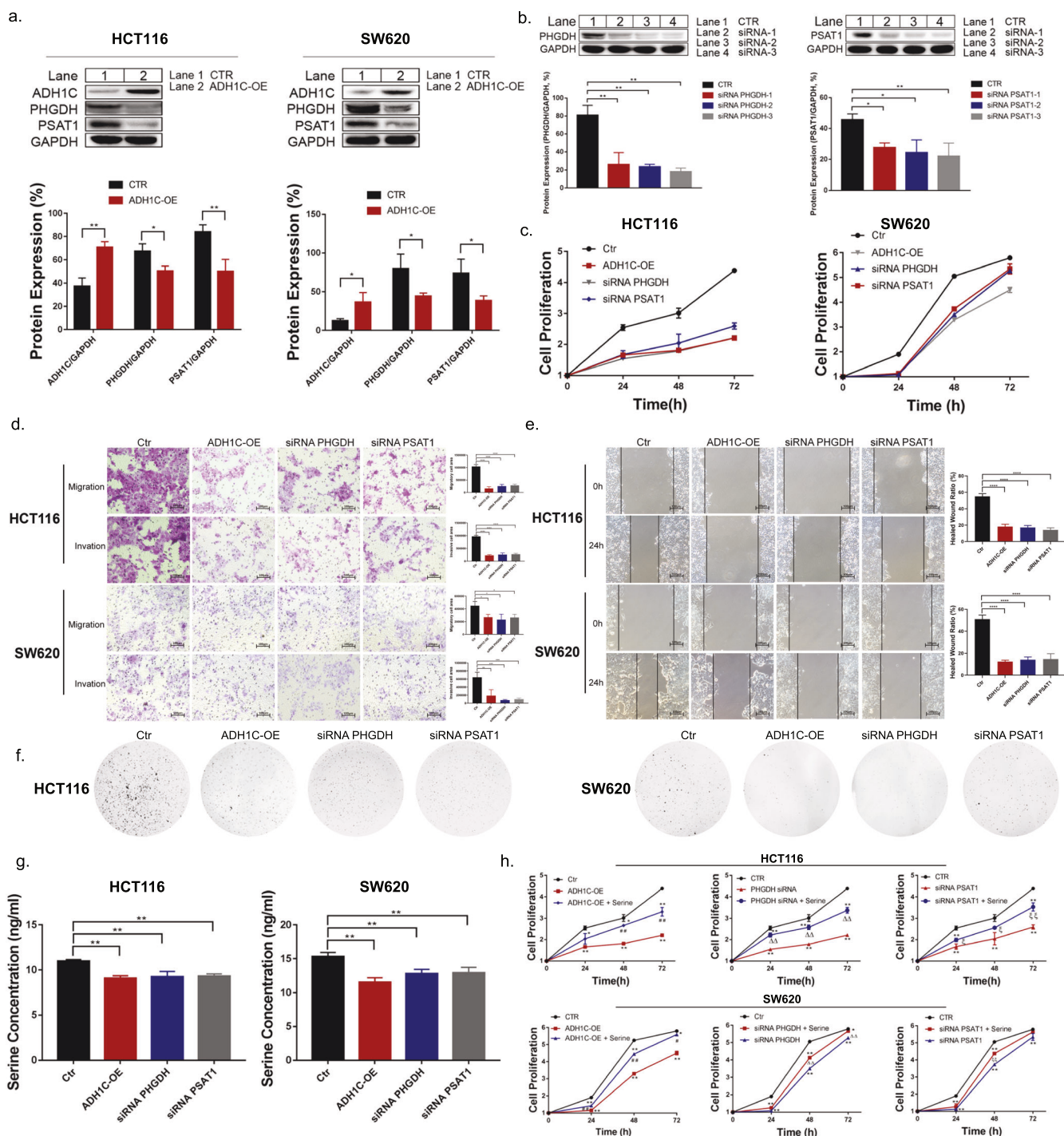


Fig. 6 ADH1C suppresses CRC via inhibiting expression of phosphatidylserine synthase PHGDH and PSAT1. **a** Overexpression of ADH1C significantly reduced expression of PHGDH and PSAT1 mRNA and protein expression in CRC cell lines. **b** Protein level verification of the knockdown effects on PHGDH and PSAT1 of several siRNA. **c–f** Overexpression of ADH1C, knockdown of PHGDH or PSAT1 expression resulted in the same phenotype in CRC cells. Overexpression of ADH1C, knockdown of PHGDH or PSAT1 inhibited cell proliferation (**c**) invasion (**d**) migration (**e**) and colony formation (**f**) $*P < 0.05$, $**P < 0.01$, $***P < 0.001$, $****P < 0.0001$ vs control. **g** Overexpression of ADH1C, knockdown of PHGDH or PSAT1 expression can significantly reduce the intracellular serine level. There is no significant difference between results caused by overexpression of ADH1C, knockdown of PHGDH or PSAT1 expression. **h** Exogenous addition of serine can reverse the effect of ADH1C overexpression, knockdown of PHGDH or PSAT1 on CRC cells. $*P < 0.05$; $**P < 0.01$ compared with control; $\#P < 0.05$, $\#\#P < 0.01$ compared with ADH1C-OE; $\Delta\Delta P < 0.01$ compared with siRNA PHGDH; $\Delta\Delta\Delta P < 0.05$; $\Delta\Delta\Delta\Delta P < 0.01$ compared with siRNA PSAT1.

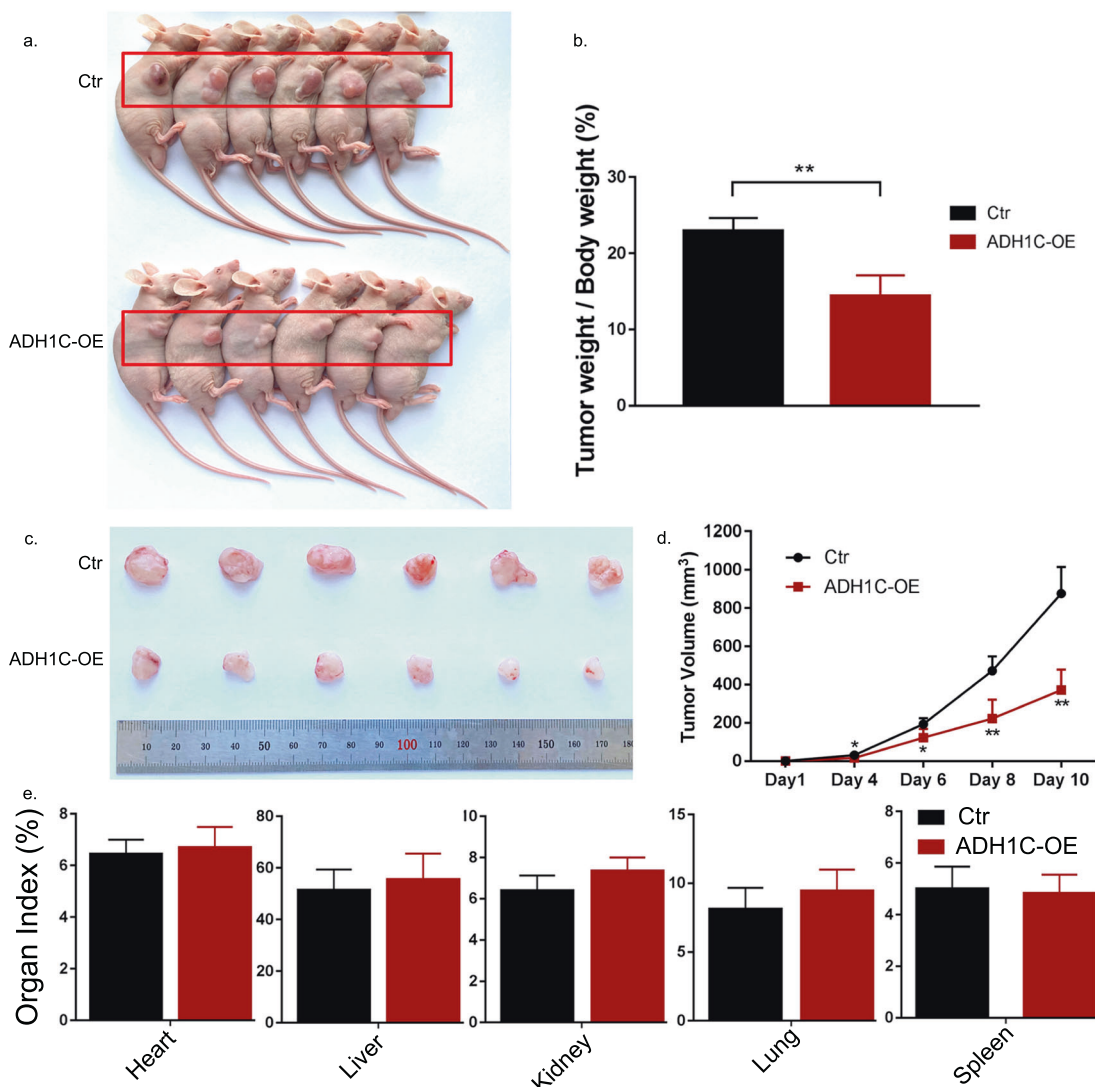


Fig. 7 ADH1C functions as a suppressor of CRC in vivo. **a** Photos of mice with xenograft tumors generated by HCT116 cells and HCT116 cells with high ADH1C expression. **b** Overexpression of ADH1C inhibited growth of xenograft CRC tumors. **c** Photos of tumors generated by HCT116 cells with or without high ADH1C expression. **d** Overexpression of ADH1C significantly reduced tumor volume of xenograft CRC tumors. **e** Overexpression of ADH1C did not bring side effect on major organs of mice. * $P < 0.05$; ** $P < 0.01$.

ADH1C has been characterized as an important member of the alcohol dehydrogenase system [33]. Ethanol is converted into acetaldehyde by ADH family members and catalyzed sequentially to acetate by ALDH [34, 35]. High expression of ADH1C was found to protect patients with non-small-cell lung cancer (NSCLC) [36] and was identified by machine learning algorithms as a promising novel biomarker for predicting lung adenocarcinoma prognosis [7], as well as hepatocellular carcinoma (HCC) [5, 6]. In our study, ADH1C was identified as a potential tumor suppressor of CRC and was associated with the overall survival of CRC using comprehensive bioinformatics methods, which further confirmed previous results. In addition, the essential role of ADH1C in inhibiting CRC was first demonstrated in our study: overexpression of ADH1C reduced the proliferation, migration, invasion and colony formation of several CRC cell lines and inhibited the growth of xenografted tumors in immunodeficient mice.

The central role of metabolic pathways and metabolites in facilitating the biosynthesis and bioenergetics required for cell proliferation and survival has been revealed over the past decade. Some metabolic reprogramming processes, such as the Warburg effect, have long been observed in cancers [37], and others, such as

alcohol, nonessential amino acid, and lipid metabolism, have also attracted increasing attention [34, 38–41]. Although ADH1C has been shown to be closely associated with several cancers [42, 43], the detailed mechanisms have not been determined, except insofar as acetaldehyde may cause genome instability [44]. In our study, advanced multiomics methods, including transcriptomics, proteomics and metabolomics, were used to explore the mechanism of ADH1C, and several potential targets of ADH1C were discovered.

We first found that overexpression of ADH1C had an inhibitory effect by reducing the expression of PHGDH and PSAT1. Abnormal expression and genetic deficiency of PHGDH have been observed in various cancers, including HCC, breast cancer, melanoma, lung cancer, glioma, colon carcinoma, leukemia, multiple myeloma, and lymphoma [45]. In human colon carcinomas, PHGDH expression was elevated tenfold and was recently identified as an independent prognostic factor of CRC patients [22, 46]. In our study, the inhibitory effect of PHGDH knockdown on CRC cells was also confirmed. Elevated PSAT1 is associated with poor clinical outcomes of patients with NSCLC, ER-negative breast cancer, ovarian cancer, nasopharyngeal carcinoma, esophageal squamous cell carcinoma and pancreatic cancer [47]. Ectopic overexpression of PSAT1 in the

CRC cell line SW480 promoted proliferation and increased tumorigenic potential in xenografted mice [23]. We also found that knockdown of PSAT1 resulted in slower proliferation of HCT116 and SW620 CRC cells and significantly reduced migration, invasion and colony formation, which further confirmed previous results.

Serine, which is a nonessential amino acid, was found to be the third most prevalent metabolite after glucose and glutamine [41] and is essential for protein, nucleotide and lipid synthesis in tumor cells, especially nucleotide synthesis. Serine provides head groups for sphingolipid and phospholipid synthesis and serves as a precursor for cellular glycine and one-carbon (1C) units [38]. Indeed, it was reported that accelerated consumption of serine was observed in cancer cells by genomic profiling [48]. An isonitrogenous diet lacking serine and glycine led to smaller tumors in mice bearing HCT116 or HT29 xenograft tumors, and the tumor size was negatively correlated with the duration of diet feeding [49]. Therefore, we believe that ADH1C reduced intracellular serine synthesis by inhibiting the important enzymes PHGDH and PSAT1 in the serine synthesis pathway. In our study, overexpression of ADH1C significantly reduced the expression of PHGDH and PSAT1 in CRC cells. Intracellular serine content was reduced both by overexpression of ADH1C and knockdown of PHGDH or PSAT1 expression, and exogenous serine restored the malignancy of CRC cells. Thus, we concluded that the mechanism of ADH1C's functional role in CRC involved inhibiting serine synthase and reducing serine biosynthesis.

There are several hypotheses that are of interest to us, which we hope to explore in future studies. For example, does ADH1C change the expression of asparagine synthetase (ASNS) and affect the metabolic processes in which ASNS is involved, as discovered

in our bioinformatics and multiomics studies? Some preliminary studies have shown that exogenous supplementation with asparagine (>100 μ M), the main metabolite derived from aspartate catalyzed by ASNS, promoted the proliferation of HCT116 cells in a glucose-deficient environment (Supplementary Fig. 9a). Supplementation with 300 μ M asparagine partially rescued the decreased cell proliferation of HCT116 cells caused by ADH1C-OE (Supplementary Fig. 9b). Does ADH1C overexpression down-regulate the ASNS protein expression in CRC, thereby reducing the biosynthesis of asparagine? Does ADH1C regulate PHGDH, PSAT1 and ASNS directly, or is the effect mediated by one or more specific transcription factors? One transcription factor, ATF4, which simultaneously regulates PHGDH, PSAT1, and ASNS and influences the biosynthesis of serine and asparagine, has attracted our attention [50–52] as a possible target of ADH1C, mediating the regulation of PHGDH and PSAT1. However, our experimental results showed that overexpression of ADH1C did not cause any significant changes in ATF4 protein expression in a variety of CRC cells (Supplementary Fig. 9c). More research is needed to answer these questions and more precisely define the mechanism by which ADH1C regulates the PHGDH/PSAT1/serine pathway.

In summary, through the comprehensive application of transcriptomics, proteomics, metabolomics and in silico analysis, followed by molecular biological experiments, ADH1C was identified as a target gene that was closely associated with the development of CRC. Overexpression of ADH1C inhibited the growth, migration, invasion and colony formation of CRC cell lines and inhibited the growth of xenografted CRC tumors in vivo. ADH1C exerted inhibition partially by reducing the expression of PHGDH or PSAT1 and serine levels in CRC cells (Fig. 8) All these

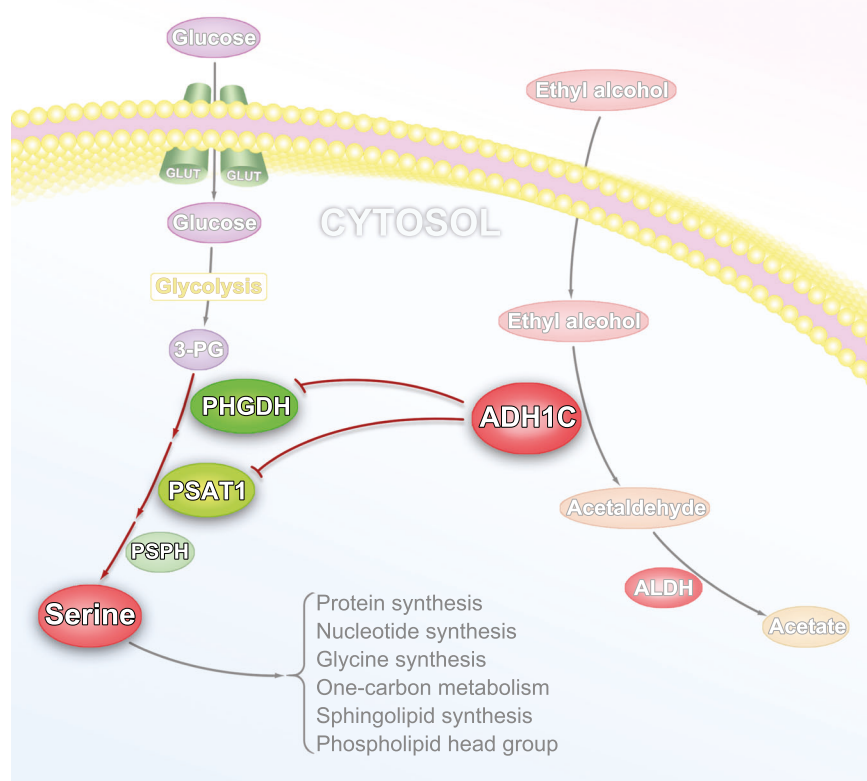


Fig. 8 ADH1C inhibits progression of CRC by ADH1C/PSAT1/serine metabolic pathway. ADH1C is a member of alcohol dehydrogenase system. Overexpression of ADH1C significantly reduced expression of PHGDH or PSAT1 mRNA and protein in CRC cell lines. Overexpression of ADH1C also reduced serine level in CRC cell lines. Knockdown of PHGDH or PSAT1 expression reduced serine level in CRC cell lines. Serine serves as precursor for cellular glycine and one-carbon metabolism, which was associated with functional roles of cancer cells. Overexpression of ADH1C inhibited proliferation, migration and invasion, colony formation of CRC cells by reducing expression of PHGDH or PSAT1, then serine level in CRC cells, which can be reversed by adding exogenous serine into culture medium. All in all, ADH1C inhibits progression of CRC by PHGDH or PSAT1/Serine metabolic pathway

results indicates that ADH1C plays an inhibitory role in the development of CRC by regulating the PHGDH and PSAT1/serine metabolic pathways and that ADH1C might be a potential drug target in CRC.

ACKNOWLEDGEMENTS

This research was funded by Beijing Natural Science Foundation (7212157), CAMS Innovation Fund for Medical Sciences (2021-1-12M-029), National Natural Science Foundation of China (81803584, 81703536), Technology Major Projects for "Major New Drugs Innovation and Development" (2018ZX09711001-005-025, 2018ZX09711001-012), National Natural Science Foundation of China (82073311).

AUTHOR CONTRIBUTIONS

JHW and GHD developed the hypothesis, designed the experiments, and revised the manuscript. SL conducted most experiments and wrote the main manuscript. HY, JYL and LWR accomplished the rest experiments, WL, YHY, BBG, YZZ, WQF and XJZ collected data and performed the statistical analysis. All authors read and approved the final manuscript.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41401-022-00894-7>.

Competing interests: The authors declare no competing interests.

REFERENCES

- Brenner H, Kloor M, Pox CP. Colorectal cancer. *Lancet*. 2014;383:1490–502. [https://doi.org/10.1016/S0140-6736\(13\)61649-9](https://doi.org/10.1016/S0140-6736(13)61649-9).
- Siegel RL, Miller KD, Goding Sauer A, Fedewa SA, Butterly LF, Anderson JC, et al. Colorectal cancer statistics, 2020. *CA: a cancer journal for clinicians*. 2020;70:145–64. <https://doi.org/10.3322/caac.21601>.
- Gu MJ, Huang QC, Bao CZ, Li YJ, Li XQ, Ye D, et al. Attributable causes of colorectal cancer in China. *BMC Cancer*. 2018;18:38 <https://doi.org/10.1186/s12885-017-3968-z>.
- Martinez-Reyes I, Chandel NS. Cancer metabolism: looking forward. *Nat Rev Cancer*. 2021;21:669–80. <https://doi.org/10.1038/s41568-021-00378-6>.
- Liu X, Li T, Kong D, You H, Kong F, Tang R. Prognostic implications of alcohol dehydrogenases in hepatocellular carcinoma. *BMC Cancer*. 2020;20:1204. <https://doi.org/10.1186/s12885-020-07689-1>.
- Gupta MK, Behara SK, Vadde R. In silico analysis of differential gene expressions in biliary stricture and hepatic carcinoma. *Gene*. 2017;597:49–58. <https://doi.org/10.1016/j.gene.2016.10.032>.
- Shen XY, Liu XP, Song CK, Wang YJ, Li S, Hu WD. Genome-wide analysis reveals alcohol dehydrogenase 1C and secreted phosphoprotein 1 for prognostic biomarkers in lung adenocarcinoma. *J Cell Physiol*. 2019;234:22311–20. <https://doi.org/10.1002/jcp.28797>.
- Liang L, Liu M, Sun X, Yuan Y, Peng K, Rashid K, et al. Identification of key genes involved in tumor immune cell infiltration and cetuximab resistance in colorectal cancer. *Cancer Cell Int*. 2021;21:135. <https://doi.org/10.1186/s12935-021-01829-8>.
- Ghosh S, Bankura B, Ghosh S, Saha ML, Pattanayak AK, Ghatak S, et al. Polymorphisms in ADH1B and ALDH2 genes associated with the increased risk of gastric cancer in West Bengal, India. *BMC Cancer*. 2017;17:782. <https://doi.org/10.1186/s12885-017-3713-7>.
- Wang L, Zhang Y, Ding D, He X, Zhu Z. Lack of association of ADH1C genotype with breast cancer susceptibility in Caucasian population: a pooled analysis of case-control studies. *Breast*. 2012;21:435–9. <https://doi.org/10.1016/j.breast.2012.01.007>.
- Wu M, Chang SC, Kampman E, Yang J, Wang XS, Gu XP, et al. Single nucleotide polymorphisms of ADH1B, ADH1C and ALDH2 genes and esophageal cancer: a population-based case-control study in China. *Int J Cancer*. 2013;132:1868–77. <https://doi.org/10.1002/ijc.27803>.
- Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, et al. NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Res*. 2013;41:D991–D5. <https://doi.org/10.1093/nar/gks1193>.
- Lee D, Park Y, Kim S Towards multi-omics characterization of tumor heterogeneity: a comprehensive review of statistical and machine learning approaches. *Brief Bioinform*. 2021;22. <https://doi.org/10.1093/bib/bbaa188>.
- Davis S, Meltzer PS. GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. *Bioinformatics*. 2007;23:1846–7.

- Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinform*. 2008;9:559. <https://doi.org/10.1186/1471-2105-9-559>.
- Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet*. 2000;25:25–9.
- Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*. 2000;28:27–30.
- Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*. 2009;4:44–57. <https://doi.org/10.1038/nprot.2008.211>.
- Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, et al. Circos: an information aesthetic for comparative genomics. *Genome Res*. 2009;19:1639–45. <https://doi.org/10.1101/gr.092759.109>.
- Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;13:2498–504.
- Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY. cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst Biol*. 2014;8:S11. <https://doi.org/10.1186/1752-0509-8-S4-S11>.
- Jia XQ, Zhang S, Zhu HJ, Wang W, Zhu JH, Wang XD, et al. Increased expression of PHGDH and prognostic significance in colorectal cancer. *Transl Oncol*. 2016;9:191–6. <https://doi.org/10.1016/j.tranon.2016.03.006>.
- Vié N, Copois V, Bascoul-Mollevis C, Denis V, Bec N, Robert B, et al. Overexpression of phosphoserine aminotransferase PSAT1 stimulates cell growth and increases chemoresistance of colon cancer cells. *Mol Cancer*. 2008;7:14. <https://doi.org/10.1186/1476-4598-7-14>.
- Deng L, Yao P, Li L, Ji F, Zhao S, Xu C, et al. p53-mediated control of aspartate-asparagine homeostasis dictates LKB1 activity and modulates cell survival. *Nat Commun*. 2020;11. <https://doi.org/10.1038/s41467-020-15573-6>.
- Li S, Liu J, Zheng X, Ren L, Yang Y, Li W, et al. Tumorigenic bacteria in colorectal cancer: mechanisms and treatments. *Cancer Biol Med*. 2021. <https://doi.org/10.20892/j.issn.2095-3941.2020.0651>.
- Matsuo K, Nishimura M, Komurov K, Shahzad MMK, Ali-Fehmi R, Roh JW, et al. Platelet-derived growth factor receptor alpha (PDGFRα) targeting and relevant biomarkers in ovarian carcinoma. *Gynecol Oncol*. 2014;132:166–75. <https://doi.org/10.1016/j.ygyno.2013.10.027>.
- Wang YZ, Qiu SC. Prediction of key genes in ovarian cancer treated with decitabine based on network strategy. *Oncol Rep*. 2016;35:3548–58. <https://doi.org/10.3892/or.2016.4697>.
- Niemira M, Collin F, Szalkowska A, Bielska A, Chwiałkowska K, Reszec J, et al. Molecular signature of subtypes of non-small-cell lung cancer by large-scale transcriptional profiling: identification of key modules and genes by weighted gene co-expression network analysis (WGCNA). *Cancers*. 2019;12. <https://doi.org/10.3390/cancers12010037>.
- Xiang Z, Li J, Song S, Wang J, Cai W, Hu W, et al. A positive feedback between IDO1 metabolite and COL12A1 via MAPK pathway to promote gastric cancer metastasis. *J Exp Clin Cancer Res*. 2019;38:314. <https://doi.org/10.1186/s13046-019-1318-5>.
- Wang Z, Tang W, Yuan J, Qiang B, Han W, Peng X. Integrated analysis of RNA-binding proteins in glioma. *Cancers*. 2020;12. <https://doi.org/10.3390/cancers12040892>.
- Yan J, Wu L, Jia C, Yu S, Lu Z, Sun Y, et al. Development of a four-gene prognostic model for pancreatic cancer based on transcriptome dysregulation. *Aging*. 2020;12:3747–70. <https://doi.org/10.18632/aging.102844>.
- Lian P, Wang Q, Zhao Y, Chen C, Sun X, Li H, et al. An eight-long non-coding RNA signature as a candidate prognostic biomarker for bladder cancer. *Aging*. 2019;11:6930–40. <https://doi.org/10.18632/aging.102225>.
- Teschke R. Alcoholic liver disease: alcohol metabolism, cascade of molecular mechanisms, cellular targets, and clinical aspects. *Biomedicines*. 2018;6. <https://doi.org/10.3390/biomedicines6040106>.
- Rodriguez FD, Coveñas R. Biochemical mechanisms associating alcohol use disorders with cancers. *Cancers*. 2021;13. <https://doi.org/10.3390/cancers13143548>.
- Crabb DW, Matsumoto M, Chang D, You M. Overview of the role of alcohol dehydrogenase and aldehyde dehydrogenase and their variants in the genesis of alcohol-related pathology. *Proc Nutr Soc*. 2004;63:49–63.
- Guo B, Zhang H, Wang J, Wu R, Zhang J, Zhang Q, et al. Identification of the signature associated with mA RNA methylation regulators and mA-related genes and construction of the risk score for prognostication in early-stage lung adenocarcinoma. *Front Genet*. 2021;12:656114. <https://doi.org/10.3389/fgene.2021.656114>.
- Hosios AM, Manning BD. Cancer signaling drives cancer metabolism: AKT and the Warburg effect. *Cancer Res*. 2021;81:4896–8. <https://doi.org/10.1158/0008-5472.CAN-21-2647>.
- Li AM, Ye J. Reprogramming of serine, glycine and one-carbon metabolism in cancer. *Biochim Biophys Acta Mol Basis Dis*. 2020;1866:165841. <https://doi.org/10.1016/j.bbdis.2020.165841>.
- Zhang J, Pavlova NN, Thompson CB. Cancer cell metabolism: the essential role of the nonessential amino acid, glutamine. *EMBO J*. 2017;36:1302–15. <https://doi.org/10.15252/emboj.201696151>.

40. DeBerardinis RJ. Serine metabolism: some tumors take the road less traveled. *Cell Metab.* 2011;14:285–6. <https://doi.org/10.1016/j.cmet.2011.08.004>.
41. Dolfi SC, Chan LLY, Qiu J, Tedeschi PM, Bertino JR, Hirshfield KM, et al. The metabolic demands of cancer cells are coupled to their size and protein synthesis rates. *Cancer Metab.* 2013;1:20. <https://doi.org/10.1186/2049-3002-1-20>.
42. Hashibe M, McKay JD, Curado MP, Oliveira JC, Koifman S, Koifman R, et al. Multiple ADH genes are associated with upper aerodigestive cancers. *Nat Genet.* 2008;40:707–9. <https://doi.org/10.1038/ng.151>.
43. Suo C, Yang Y, Yuan Z, Zhang T, Yang X, Qing T, et al. Alcohol intake interacts with functional genetic polymorphisms of aldehyde dehydrogenase (ALDH2) and alcohol dehydrogenase (ADH) to increase esophageal squamous cell cancer risk. *J Thorac Oncol.* 2019;14:712–25. <https://doi.org/10.1016/j.jtho.2018.12.023>.
44. Tan SLW, Chadha S, Liu Y, Gabasova E, Perera D, Ahmed K, et al. A class of environmental and endogenous toxins induces BRCA2 haploinsufficiency and genome instability. *Cell.* 2017;169. <https://doi.org/10.1016/j.cell.2017.05.010>.
45. Zhao JY, Feng KR, Wang F, Zhang JW, Cheng JF, Lin GQ, et al. A retrospective overview of PHGDH and its inhibitors for regulating cancer metabolism. *Eur J Med Chem.* 2021;217:113379. <https://doi.org/10.1016/j.ejmech.2021.113379>.
46. Snell K, Natsumeda Y, Eble JN, Glover JL, Weber G. Enzymic imbalance in serine metabolism in human colon carcinoma and rat sarcoma. *Br J Cancer.* 1988;57:87–90.
47. Biyik-Sit R, Kruer T, Dougherty S, Bradley JA, Wilkey DW, Merchant ML, et al. Nuclear pyruvate kinase M2 (PKM2) contributes to phosphoserine aminotransferase 1 (PSAT1)-mediated cell migration in egfr-activated lung cancer cells. *Cancers.* 2021;13. <https://doi.org/10.3390/cancers13163938>.
48. Possemato R, Marks KM, Shaul YD, Pacold ME, Kim D, Birsoy K, et al. Functional genomics reveal that the serine synthesis pathway is essential in breast cancer. *Nature.* 2011;476:346–50. <https://doi.org/10.1038/nature10350>.
49. Muthusamy T, Cordes T, Handzlik MK, You L, Lim EW, Gengatharan J, et al. Serine restriction alters sphingolipid diversity to constrain tumour growth. *Nature.* 2020;586:790–5. <https://doi.org/10.1038/s41586-020-2609-x>.
50. Williams RT, Guarecuco R, Gates LA, Barrows D, Passarelli MC, Carey B, et al. ZBTB1 regulates asparagine synthesis and leukemia cell response to *L*-asparaginase. *Cell Metab.* 2020;31:852–61.e6. <https://doi.org/10.1016/j.cmet.2020.03.008>.
51. Ye J, Mancuso A, Tong X, Ward PS, Fan J, Rabinowitz JD, et al. Pyruvate kinase M2 promotes de novo serine synthesis to sustain mTORC1 activity and cell proliferation. *Proc Natl Acad Sci USA.* 2012;109:6904–9. <https://doi.org/10.1073/pnas.1204176109>.
52. Ye J, Kumanova M, Hart LS, Sloane K, Zhang H, De Panis DN, et al. The GCN2-ATF4 pathway is critical for tumour cell survival and proliferation in response to nutrient deprivation. *EMBO J.* 2010;29:2082–96. <https://doi.org/10.1038/emboj.2010.81>.