



## OPEN ACCESS

## EDITED BY

Francesca Negri,  
University Hospital of Parma, Italy

## REVIEWED BY

Albrecht Reichle,  
University Medical Center Regensburg,  
Germany  
Yu-Chan Chang,  
National Yang Ming Chiao Tung  
University, Taiwan

## \*CORRESPONDENCE

Yongwen He  
heyongwen2@sina.com  
Zhenhui Li  
lizhenhui621@qq.com  
Tao Zhang  
taozhang@sdu.edu.cn

<sup>†</sup>These authors have contributed  
equally to this work and share  
first authorship

## SPECIALTY SECTION

This article was submitted to  
Gastrointestinal Cancers:  
Colorectal Cancer,  
a section of the journal  
Frontiers in Oncology

RECEIVED 09 May 2022

ACCEPTED 11 July 2022

PUBLISHED 01 August 2022

## CITATION

You D, Zhang S, Yan S, Ding Y, Li C,  
Cheng X, Wu L, Wang W, Zhang T, Li Z  
and He Y (2022) SAMHD1 as a  
prognostic and predictive biomarker in  
stage II colorectal cancer: A  
multicenter cohort study.  
*Front. Oncol.* 12:939982.  
doi: 10.3389/fonc.2022.939982

## COPYRIGHT

© 2022 You, Zhang, Yan, Ding, Li,  
Cheng, Wu, Wang, Zhang, Li and He.  
This is an open-access article  
distributed under the terms of the  
Creative Commons Attribution License  
(CC BY). The use, distribution or  
reproduction in other forums is  
permitted, provided the original  
author(s) and the copyright owner(s)  
are credited and that the original  
publication in this journal is cited, in  
accordance with accepted academic  
practice. No use, distribution or  
reproduction is permitted which does  
not comply with these terms.

# SAMHD1 as a prognostic and predictive biomarker in stage II colorectal cancer: A multicenter cohort study

Dingyun You<sup>1,2,3†</sup>, Shuai Zhang<sup>4†</sup>, Shan Yan<sup>3†</sup>, Yingying Ding<sup>5†</sup>,  
Chunxia Li<sup>4</sup>, Xianshuo Cheng<sup>6</sup>, Lin Wu<sup>7</sup>, Weizhou Wang<sup>8</sup>,  
Tao Zhang<sup>4\*</sup>, Zhenhui Li<sup>5\*</sup> and Yongwen He<sup>1\*</sup>

<sup>1</sup>Department of Dental Research, The Affiliated Stomatological Hospital of Kunming Medical University, Kunming, China, <sup>2</sup>Yunnan Key Laboratory of Stomatology, Kunming Medical University, Kunming, China, <sup>3</sup>Yunnan Key Laboratory of Stem Cell and Regenerative Medicine, Biomedical Engineering Research Center, Kunming Medical University, Kunming, China, <sup>4</sup>Department of Biostatistics, School of Public Health, Cheeloo College of Medicine, Shandong University, Jinan, China, <sup>5</sup>Department of Radiology, The Third Affiliated Hospital of Kunming Medical University, Yunnan Cancer Hospital, Yunnan Cancer Center, Kunming, China, <sup>6</sup>Department of Colorectal Surgery, The Third Affiliated Hospital of Kunming Medical University, Yunnan Cancer Hospital, Yunnan Cancer Center, Kunming, China, <sup>7</sup>Department of Pathology, The Third Affiliated Hospital of Kunming Medical University, Yunnan Cancer Hospital, Yunnan Cancer Center, Kunming, China, <sup>8</sup>Department of Orthopedics, The First Affiliated Hospital of Kunming Medical University, Kunming, China

**Background:** The identification of high-risk population patients is key to the personalized treatment options for the stage II colorectal cancers. The use of proteomics in the prognosis of patients with stage II colorectal cancer remains unclear.

**Methods:** Using quantitative proteomics, we analyzed proteins that are differentially expressed in the tumor and adjacent normal tissues of 11 paired colorectal cancer patients with and without recurrence selected by a nested case-control design. Of the 21 identified proteins, we selected one candidate protein. The association of the corresponding gene of the selected protein with overall survival (OS) and adjuvant chemotherapy was analyzed using two independent cohorts of patients with stages II colorectal cancer.

**Results:** Sterile  $\alpha$  motif and histidine-aspartate domain-containing protein 1 (SAMHD1) was selected as the candidate biomarker. A group of 124 patients (12.5%) were stratified into SAMHD1-high subgroup. The 5-year OS rate of SAMHD1-high patients was lower than that of SAMHD1-low patients with stage II colorectal cancer (discovery cohort: hazard ratio [HR] = 2.89, 95% confidence interval [CI], 1.17-7.18,  $P = 0.016$ ; validation cohort: HR = 2.25, 95% CI, 1.17-4.34,  $P = 0.013$ ). The Cox multivariate analysis yielded similar results. In a pooled database, the 5-year OS rate was significantly different between patients with and without adjuvant chemotherapy among stage II SAMHD1-low tumors than in patients with stage II SAMHD1-high tumors (88% vs. 77%,  $P = 0.032$ ).

**Conclusions:** SAMHD1-high expression could help in identifying patients with stage II colorectal cancer with poor prognosis and less benefit from adjuvant chemotherapy.

#### KEYWORDS

SAMHD1, colorectal cancer, Cox model, prognostic markers, nested case-control design, MSI

## Introduction

Globally, colorectal cancer is the third most common malignant tumor and the second leading cause of cancer-related deaths (1). In China, colorectal cancer poses a huge health burden, with more than 290,000 deaths reported annually (2). Local recurrence and distant metastasis are the major reasons for the high mortality rate in patients with resectable colorectal cancer (3). The rate of recurrence for stages II and III colorectal cancer is approximately 20% and 48%, respectively (4). To reduce the incidence of recurrence, adjuvant chemotherapy following total meso-rectal excision is the standard of care for stage III patients according to international guidelines (5–8); wherein patients with stage III colorectal cancer who received adjuvant chemotherapy showed significant improvement in survival (9). However, patients with stage II colorectal cancer showed minimal improvement in the 5-year overall survival (OS) rate (2%–5%) (10). There is no consensus on whether patients with stage II colorectal cancer could benefit from adjuvant chemotherapy; therefore, recommending adjuvant chemotherapy for those patients is still controversial (5–8). Therefore, it is crucial to identify patients with stage II colorectal cancer who could benefit from adjuvant chemotherapy.

Prognostic risk factors are essential to help clinicians make better-informed decisions while selecting the best treatment strategy for patients with stage II colorectal cancer and determining the need for adjuvant treatment (11). At present, the major well-known prognostic risk factors for patients with colorectal cancer are stage pT4, bowel perforation or occlusion, lymphatic-vascular-perineural invasion, poorly differentiated histology (excluding microsatellite instability-high [MSI-H] tumors), inadequate lymph node sampling, and positive margins after surgery (5–8). However, all these factors, except stage pT4, are insufficient to identify patients with stage II colorectal cancer who could benefit from adjuvant chemotherapy (12). In the last few years, several efforts have been made to identify novel biomarkers that are able to predict a higher risk of relapse in patients with stage II colorectal cancer, such as identifying their gene expression signatures (Oncotype,

ColoPrint, ColDX) (13–15), microRNA signatures (16), circulating tumor DNA (17–21), immune-related signatures (22–24), and deep learning signatures (25). However, high costs or complexity in the techniques of these approaches have prevented their successful translation into routine clinical practice. This has led to the emerging need for the identification of novel and more feasible biomarkers.

Over the past few years, mass-spectrometry-based proteomics has emerged as the method of choice for identifying possible prognostic indicators of outcome and disease response to therapy (26–29). We used proteomics to identify and select sterile  $\alpha$  motif and histidine-aspartate domain-containing protein 1 (SAMHD1) as the candidate biomarker based on literature reviews and experiments. Using subgroup analysis involving retrospective patient cohorts, we evaluated the association between the SAMHD1 biomarker and the benefits from adjuvant chemotherapy and survival in patients with stage II and III colorectal cancer.

## Methods

### Patients and samples

The study protocol was approved by the Yunnan Cancer Hospital Ethics Committee (No. KY2019141). The requirement for informed consent was waived by the ethics committee owing to the retrospective nature of the study. The data were anonymized. For proteomic analysis, surgically resected biopsies of patients with colorectal cancer and paired non-cancerous tissues (collected 10 cm from the tumor) were collected from 11 pairs of patients with stage II and III colorectal cancer with and without recurrence, from Yunnan Cancer Hospital. These 11 pairs of patients were selected by propensity score matching (PSM) from the original cohort, including consecutive patients with stage I–III colorectal cancer who underwent radical resection at Yunnan Cancer Hospital between December 2010 and February 2019 (referred to as the Yunnan colorectal cancer cohort). The association between the expression levels of SAMHD1 messenger RNA (mRNA) and OS was tested in a discovery dataset of 335

patients from The Cancer Genome Atlas (TCGA) and a validation dataset of 465 patients from the National Center for Biotechnology Information Gene Expression Omnibus (NCBI-GEO). Patients who received neoadjuvant treatment were excluded from the analysis. The flowchart of the study is shown in Figure 1.

### Propensity score matching

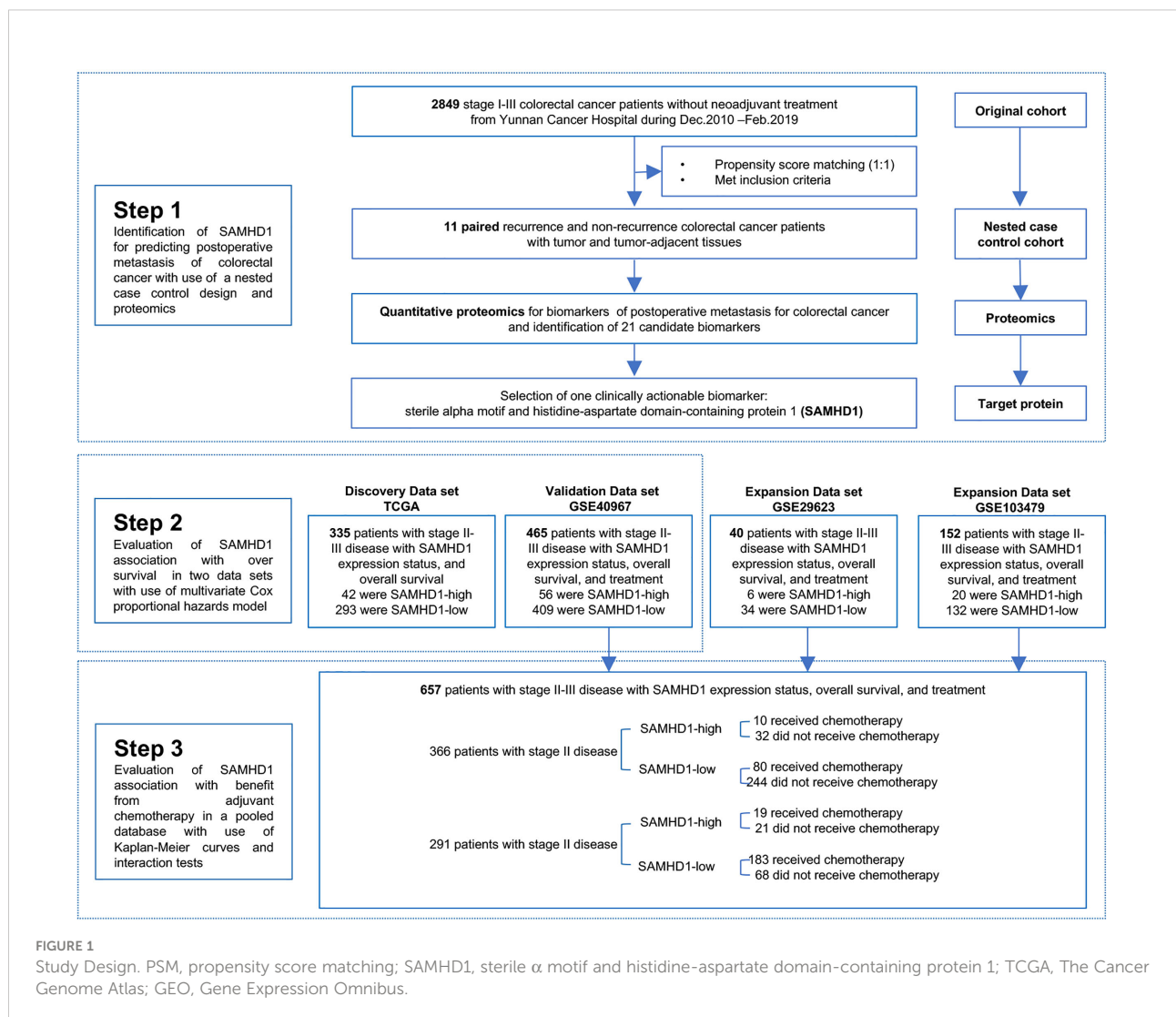
We performed PSM (30, 31) to strictly balance the critical variables between postoperative and non-postoperative metastatic patients within 3 years after surgery in the Yunnan colorectal cancer cohort. Propensity scores were generated using a logistic regression model with age, sex, body mass index, surgical pathological type, site of primary carcinoma, and pathological stage as the independent variables. Each metastatic patient was matched 1:1 to two patients in the non-

metastasis group using a 0.001 caliper width (propensity scores must be within 0.1% of each other to create a match), and the resulting matches were used in the following selection.

Subsequently, we selected matched patients according to the following inclusion criteria: (1) patients with stage II or III colorectal cancer; (2) available formalin-fixed paraffin-embedded (FFPE) specimens; (3) available data on recurrence-free survival (RFS) and OS; (4) patients without recurrence whose duration of OS is longer than that of patients with recurrence; and (5) data including cancerous and paired non-cancerous tissues. Tandem mass tag (TMT)-labelled quantitative proteomics was performed on the matched patients.

### TMT-labelling quantitative proteomics

For each patient, quantitative proteomics was performed on the tumor and tumor-adjacent tissues, and the protein was



extracted using the FFPE Total Protein Extraction Kit (Sangon Biotech, NO. C500058, Shanghai, China), according to the manufacturer's instructions. The extracted proteins were quantified using a BCA protein assay kit (Bio-Rad, USA). Protein digestion was performed according to the FASP procedure described by Wisniewski et al. (32), and the resulting peptide mixture was labeled using the 6-plex TMT reagent according to the manufacturer's instructions (Thermo Fisher Scientific, Waltham, USA). Detailed procedures for TMT labeling, peptide fractionation, and LC-MS/MS analysis are described in the [Supplementary material A](#).

## Identification of the target protein

The differentially expressed proteins between tumor and tumor-adjacent tissues were identified using the Student's *t*-test ( $P < 0.05$ ). Proteins associated with metastasis were verified using the univariate Cox regression analysis ( $P < 0.01$ ). Subsequently, we focused on the intersection of the differentially expressed and metastasis-related proteins. Proteins that had rarely been reported in most cancers, according to the literature search and our basic research, were selected for further analysis.

## Analysis of tissue microarrays in the discovery and validation datasets

Gene expression profiles for colorectal cancer tissues, fully annotated with clinical and pathological information, were obtained from two independent sources; TCGA ([Figure S1](#)) and NCBI-GEO, including GSE40967 ([Figure S2](#)). A detailed description of the patient cohorts represented by the two independent sources is provided in [Table S1](#).

Due to the considerable variation in the coverage of the sequencing platforms, pipelines, assays, and tools/algorithms between the TCGA and GEO datasets, the frequency of the identified variants was impacted (33). Taking these constraints into consideration, we used the Z-score (34) to standardize data across different experiments and to normalize the expression data of SAMHD1 from these two datasets prior to data analysis.

Subsequently, SAMHD1 expression levels were stratified into SAMHD1-high and SAMHD1-low subgroups according to the SAMHD1 expression, the threshold of which was identified in patients with stage II colorectal cancer using X-tile from the discovery dataset (35). We explored the association between the expression levels of SAMHD1, the OS outcomes, and the interaction between SAMHD1 expression level and adjuvant chemotherapy in stage II and stage III colorectal cancer.

## SAMHD1 expression and benefit from adjuvant chemotherapy

To evaluate whether patients with SAMHD1-high tumors could benefit from adjuvant chemotherapy, we investigated the association between SAMHD1 status (assessed at the mRNA level) and OS among patients who either did or did not receive adjuvant chemotherapy in the NCBI-GEO dataset by pooling the following three datasets: GSE40967, GSE29623 and GSE103479. The three datasets were found to satisfy our criteria (i.e., knowledge of pathological stage, available information on SAMHD1 expression, adjuvant chemotherapy, duration of OS, and follow-up duration) ([Figures S3, S4, and Table S1](#)).

## Statistical analysis

We downloaded the transcriptome profiles in FPKM format and the corresponding clinical information from the TCGA portal (<https://portal.gdc.cancer.gov/>) and NCBI-GEO dataset (<https://www.ncbi.nlm.nih.gov/geo/>). The NCBI-GEO datasets recruited for multiple dataset analysis and were based on different platforms. Therefore, we combined the three datasets to expand the sample size and avoid generating less reliable results by normalization using the robust multi-chip average (RMA) algorithm and removed the batch effect using the *affy* and *sva* R packages. Probes corresponding to the same gene were averaged.

Patient subgroups were compared with respect to survival outcomes using Kaplan–Meier curves, log-rank tests, and multivariate analyses based on the Cox proportional hazards method adjusting for patient age, sex, and adjuvant chemotherapy.

All analyses were conducted using R software (version 3.6.3; <http://www.R-project.org>). Statistical significance was set at  $P < 0.05$ .

## Results

### The clinical characteristics of 11 paired-patients

Eleven pairs of patients were selected through PSM. The clinical and pathological characteristics of patients with stage II or III colorectal cancer are shown in [Table S2](#). The age of the patients in the non-metastasis group ranged between 46 and 74 years. In the non-metastatic group, 5 patients were males, 8 patients were in stage II, one patient died, and no recurrence occurred. The median OS and RFS follow-up times were both 51.6 months. The age of patients in the metastatic group ranged between 42 and 75 years. In the metastatic group, five patients

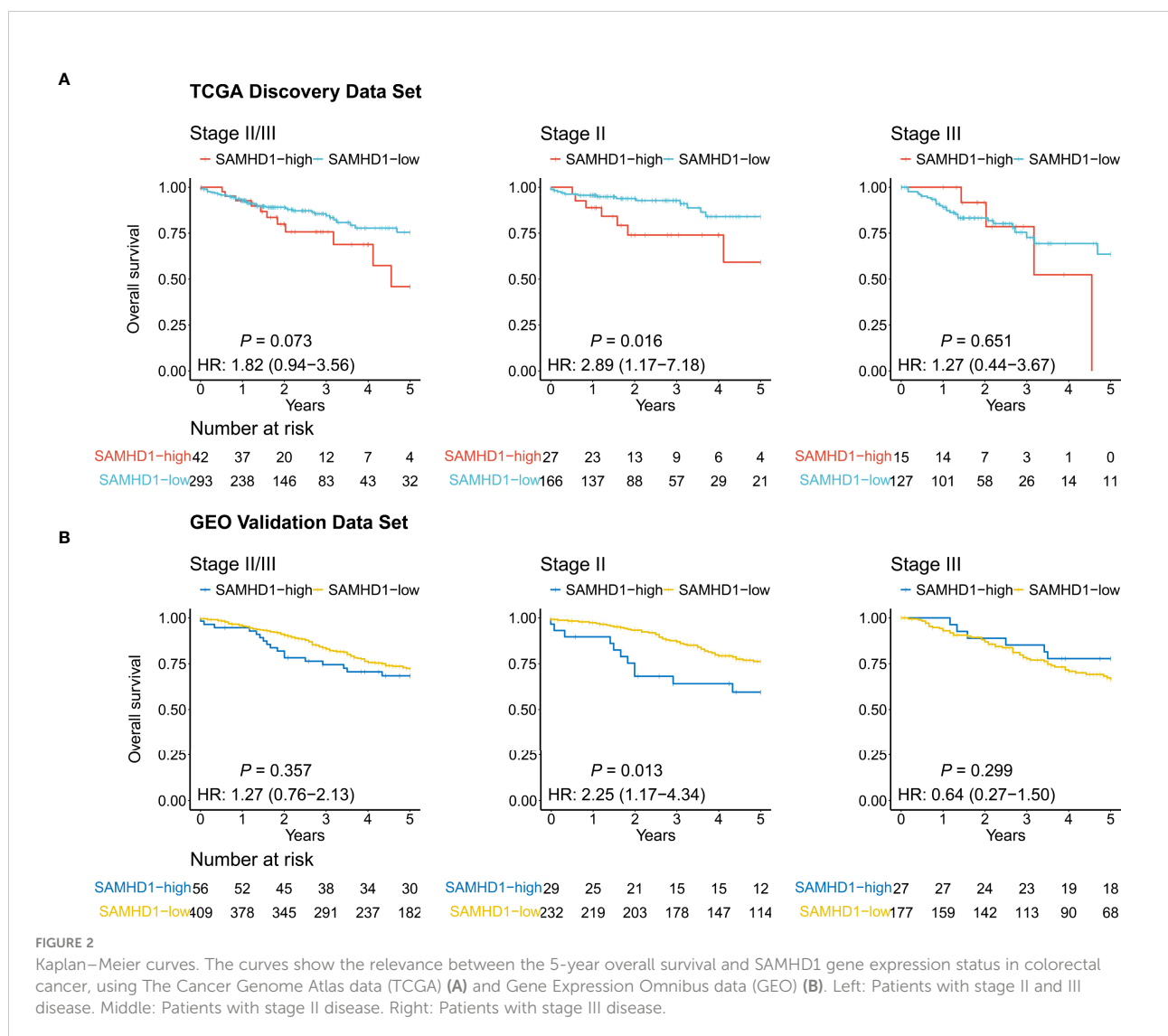
were males, eight patients had stage II disease, five patients died, and recurrence occurred in five patients. The median OS and RFS follow-up times were 31.8 and 21.4 months, respectively. Based on the univariate analysis, there was a significant difference in the OS follow-up time (51.6 months vs 31.8 months, respectively,  $P = 0.002$ ), RFS event (100% vs 54.5%, respectively,  $P = 0.042$ ), and RFS follow-up time (51.6 months vs 21.4 months, respectively,  $P < 0.001$ ) between the non-metastatic and metastatic groups. Figure S5 shows the OS and RFS curves for all patients.

### Identification of SAMHD1

A total of 5,197 proteins were identified using TMT-labelled quantitative proteomics. We processed the protein expression data by deleting proteins in which more than 50% of the samples

had missing values. Among the remaining 2,760 proteins, which were retained for further analysis, 1,409 proteins showed significance with the  $P < 0.05$  t-test threshold (Table S3); 38 proteins were associated with metastasis ( $P < 0.01$ ) in the univariate Cox regression analysis (Table S4). The volcano plot shows the distribution of  $P$ -values of the t-test and the univariate Cox regression analysis (Figure S6A). In our study, 28 candidate proteins were found to be common between the 1,409 differentially expressed proteins and the 38 metastasis-related proteins (Figure S6B). Of the 28 candidate proteins, 7 proteins were not annotated with coding genes. The information regarding the remaining 21 proteins is shown in Table S5.

Based on previous literature reviews and basic experiments, we screened these 21 proteins and finally yielded the protein SAMHD1 (36–39). Figure S7 shows the different distributions of SAMHD1 expression in the cancer tissues between the non-metastatic and metastatic groups.



The Pearson's correlation coefficients between the expression of the SAMHD1 gene and the genes associated with microsatellite instability (MLH1, MSH2, MSH6, and PMS2) were -0.012, 0.17, 0.19, and 0.28, the results indicated there was a weak correlation between SAMHD1 expression and microsatellite instability-related genes (Figure S8), and SAMHD1 has a good complementary effect with those genes. Additionally, we compared and contrasted SAMHD1 expression according to KRAS mutation status, BRAF mutation status, tumor location, and defective DNA mismatch repair status, as they were frequently mutated genes or risk parameters in colorectal cancer. Statistical significance was detected using t-test for comparisons between all those groups. The results showed that SAMHD1 gene expression only partially overlapped with tumors defined by those factors (Figure S9).

## SAMHD1 expression and OS in the discovery dataset

The optimum cutoff score for SAMHD1 expression generated by the X-tile plot was 1.15 (Figure S10). In total, the 335 patients were stratified into SAMHD1-low (n = 293, 87.5%) and SAMHD1-high (n = 42, 12.5%) groups, according to the expression of SAMHD1 in the discovery dataset. The baseline characteristics and known molecules of stage II and III in the discovery dataset were shown in Table S6.

Kaplan–Meier curves were used to compare the 5-year OS of the two groups. As shown in Figure 2A, the expression level of SAMHD1 tended to be associated with the 5-year OS (SAMHD1-high vs. SAMHD1-low; HR = 1.82; 95% CI, [0.94–3.56]; and  $P = 0.073$ ) among patients with stage II and III colorectal cancer. With respect to stage II, the 5-year OS rate of the 27 patients (13.99%) with SAMHD1-high expression level was higher than that of the 166 patients (86.01%) with

SAMHD1-low expression level (HR = 2.89; 95% CI, [1.17–7.18]; and  $P = 0.016$ ). However, there was no significant difference in the 5-year OS between the SAMHD1-low (n = 127) and SAMHD1-high (n = 15) groups with stage III colorectal cancer (HR = 1.27; 95% CI, [0.44–3.67]; and  $P = 0.651$ ). In the multivariate analysis, with adjustment of age and sex as confounding variables, the HR for OS among stage II patients with SAMHD1-high versus SAMHD1-low was 2.99 (95% CI, [1.17–7.65]; and  $P = 0.023$ ) (Table 1).

## SAMHD1 expression and OS in the validation dataset

To evaluate the robustness of our findings, we performed an analysis in the validation dataset including 56 SAMHD1-high patients (12.04%) and 409 SAMHD1-low patients (87.96%). The baseline characteristics and known molecules of stage II and III in the validation dataset were described in Table S7. As shown in Figure 2B, we observed that the high expression of SAMHD1 (n = 29) was associated with a lower 5-year OS rate than a low expression of SAMHD1 (n = 232) among stage II patients (HR = 2.25; 95% CI, [1.17–4.34]; and  $P = 0.013$ ), but not in stage III patients (n = 204; HR = 0.64; 95% CI, [0.27–1.50]; and  $P = 0.299$ ) (Table 1). After adjusting for sex, age, and adjuvant chemotherapy, multivariate analysis also confirmed that high SAMHD1 expression status was associated with shorter OS in stage II patients (HR = 2.81; 95% CI, [1.43–5.50]; and  $P = 0.003$ ) (Table 2).

The SAMHD1 expression groups had similar hazard ratios among stage II patients compared with the classical risk parameter such as age and T stage in the multivariate Cox regression analyses (Figure S11), which is based on analyses about the relative importance of each risk parameter for OS using the  $\chi^2$  proportion test in stage II patients.

TABLE 1 Univariate and multivariable Cox analyses for overall survival among patients in the discovery data set.

Subgroup	Variable	Univariate Analysis		Multivariate Analysis	
		HR (95% CI)	P value	HR (95% CI)	P value
Stage II/III	SAMHD1_high vs. SAMHD1_low	1.82 (0.94-3.56)	0.078	1.55 (0.77-3.09)	0.217
	Age <sup>a</sup>	1.04 (1.01-1.07)	0.002	1.05 (1.02-1.07)	0.001
	Male vs. Female	0.79 (0.46-1.37)	0.409	0.75 (0.43-1.32)	0.327
	Stage III vs. Stage II	2.20 (1.25-3.85)	0.006	2.58 (1.47-4.55)	0.001
Stage II	SAMHD1_high vs. SAMHD1_low	2.89 (1.17-7.18)	0.022	2.99 (1.17-7.65)	0.023
	Age <sup>a</sup>	1.12 (1.05-1.18)	<0.001	1.12 (1.05-1.20)	<0.001
	Male vs. Female	1.09 (0.46-2.58)	0.849	1.20 (0.49-2.94)	0.686
Stage III	SAMHD1_high vs. SAMHD1_low	1.27 (0.44-3.67)	0.653	0.92 (0.30-2.80)	0.888
	Age <sup>a</sup>	1.03 (1.00-1.06)	0.086	1.03 (1.00-1.06)	0.055
	Male vs. Female	0.70 (0.34-1.45)	0.334	0.59 (0.28-1.27)	0.180

<sup>a</sup>continuous variable.

CI, confidence interval; HR, hazard ratio; SAMHD1, sterile alpha motif and histidine-aspartate domain-containing protein 1.

TABLE 2 Univariate and multivariable Cox analyses for overall survival among patients in the validation data set.

Subgroup	Variable	Univariate Analysis		Multivariate Analysis	
		HR (95% CI)	P value	HR (95% CI)	P value
Stage II/III	SAMHD1_high vs. SAMHD1_low	1.27 (0.76-2.13)	0.359	1.14 (0.66-1.95)	0.634
	Age <sup>a</sup>	1.03 (1.02-1.05)	<0.001	1.03 (1.01-1.05)	0.001
	Male vs. Female	1.28 (0.88-1.86)	0.195	1.47 (1.01-2.16)	0.046
	Stage III vs. Stage II	1.30 (0.90-1.88)	0.156	1.70 (1.08-2.67)	0.021
	Adjuvant Chemotherapy <sup>b</sup>	0.71 (0.49-1.03)	0.073	0.64 (0.40-1.04)	0.071
Stage II	SAMHD1_high vs. SAMHD1_low	2.25 (1.17-4.34)	0.015	2.81 (1.43-5.50)	0.003
	Age <sup>a</sup>	1.03 (1.01-1.05)	0.016	1.04 (1.01-1.06)	0.006
	Male vs. Female	1.20 (0.71-2.04)	0.496	1.32 (0.78-2.25)	0.301
	Adjuvant Chemotherapy <sup>b</sup>	0.76 (0.40-1.47)	0.414	1.07 (0.53-2.16)	0.850
Stage III	SAMHD1_high vs. SAMHD1_low	0.64 (0.27-1.50)	0.304	0.43 (0.18-1.02)	0.056
	Age <sup>a</sup>	1.03 (1.01-1.06)	0.002	1.03 (1.01-1.05)	0.017
	Male vs. Female	1.43 (0.84-2.44)	0.185	1.60 (0.93-2.75)	0.088
	Adjuvant Chemotherapy <sup>b</sup>	0.40 (0.24-0.68)	0.001	0.40 (0.22-0.71)	0.002

<sup>a</sup>continuous variable. <sup>b</sup>yes vs. no.

CI, confidence interval; HR, hazard ratio; SAMHD1, sterile alpha motif and histidine-aspartate domain-containing protein 1.

## SAMHD1 expression and benefit from adjuvant chemotherapy

Figure 3 shows the relationship between the expression levels of SAMHD1 and benefit from adjuvant chemotherapy in 657 patients with stage II or III colorectal cancer. In the SAMHD1-low patient population, treatment with adjuvant chemotherapy was associated with higher OS in the stage II subgroup (88% with chemotherapy vs. 77% with no chemotherapy; HR = 0.49; 95% CI, [0.25–0.95], and  $P = 0.032$ ) and in the stage III subgroup (73% with chemotherapy vs. 44% with no chemotherapy; HR = 0.34; 95% CI, [0.22–0.51], and  $P < 0.001$ ) (Figure 3). In the SAMHD1-high patient population, treatment with adjuvant chemotherapy was not associated with higher OS in either the stage II subgroup (chemotherapy vs. no chemotherapy; HR = 0.67; 95% CI, [0.19–2.35], and  $P = 0.523$ ) or the stage III subgroup (chemotherapy vs. no chemotherapy, HR = 0.50; 95% CI, [0.12–1.99], and  $P = 0.312$ ) (Figure 3).

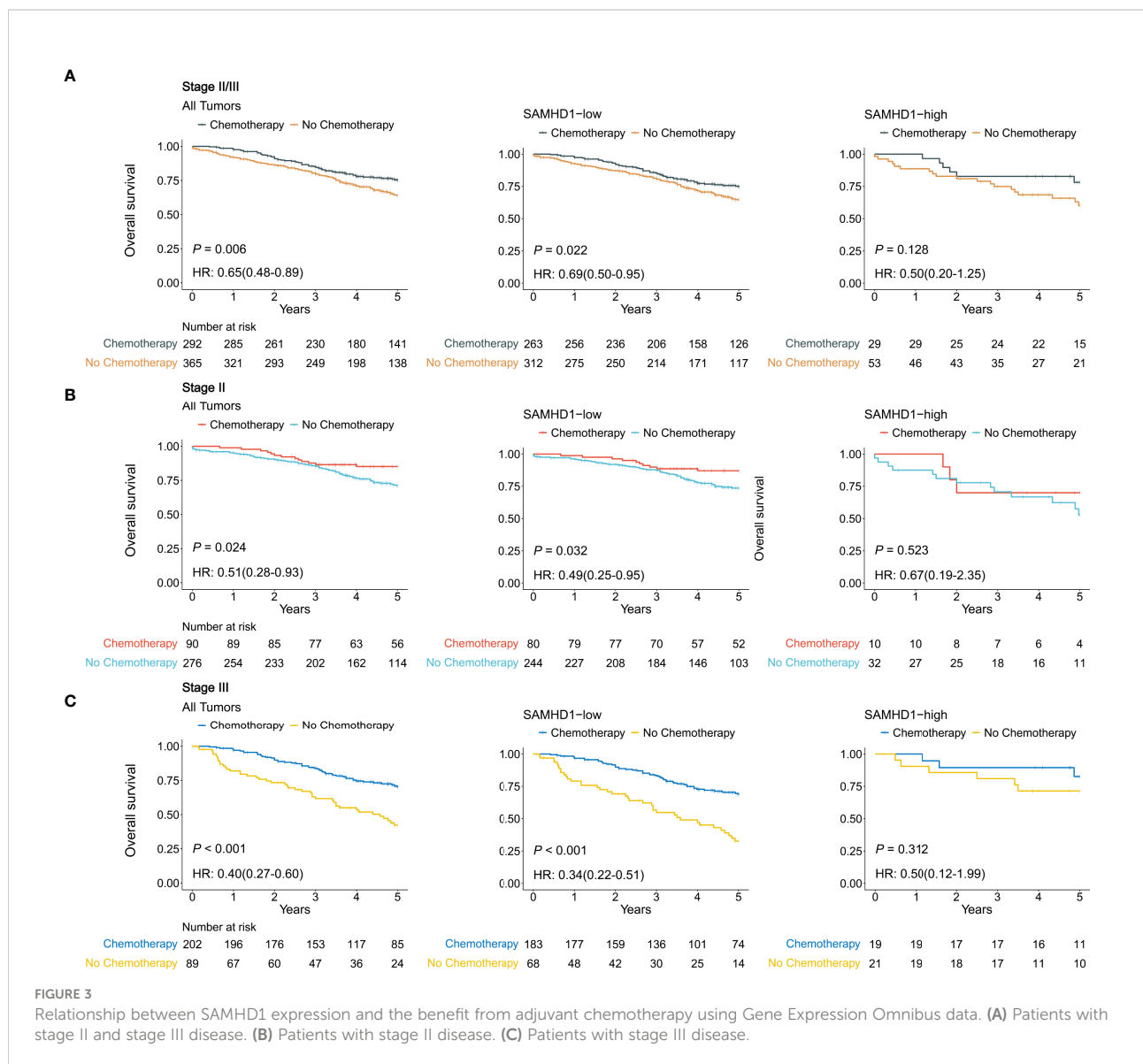
## Discussion

Using proteomics analysis, SAMHD1 was identified as a potential biomarker displaying a significant prognostic value. It was differentially expressed in the paired colorectal cancer groups with and without recurrence as selected by a nested case-control design from a large retrospective cohort. Using public colorectal cancer datasets for biomarker discovery, we illustrated that SAMHD1 had prognostic and predictive powers that could be helpful for patients with stage II colorectal cancer and had a predictive power in those with stage III colorectal cancer. Validation was performed using tissue microarrays on different cohorts of patients. Hence, SAMHD1 could complement MSI/MMR status as a molecular marker involved in the high-risk

definition for patients with stage II colorectal cancer and help in making clinical decisions for adjuvant chemotherapy for patients with stages II and III colorectal cancer.

This study showed that the high expression of SAMHD1 in stage II colorectal cancer tissues was correlated with poor prognosis. We speculated that a higher rate of mutations may occur in patients with high expression of SAMHD1 resulting in disease progression because mutations in SAMHD1 that alter its dNTPase activity are associated with colon cancer (40). Moreover, a previous study reported that SAMHD1 upregulation was found in the colorectal cancer tissue of the patients with advanced colorectal cancer compared to their normal counterparts (41). Additionally, the role of SAMHD1 in numerous types of cancer, such as chronic lymphocytic leukemia, lung cancer, and colorectal cancer, has been extensively studied (42). Moreover, the high expression level of SAMHD1 had an independent significant association with unfavorable OS in some types of cancer (37, 43, 44). Hence, the expression level of SAMHD1 could be a prognostic biomarker for stage II colorectal cancer.

This study is the first to demonstrate that SAMHD1 is a predictive biomarker for adjuvant chemotherapy in patients with stage II and III colorectal cancer. Several studies have reported that high expression of SAMHD1 negatively impacts the efficacy of nucleoside-based chemotherapies in different cohorts of patients with leukemia (36, 37, 43, 45–47). The negative role of SAMHD1 in the sensitivity to chemotherapy can be attributed to various reasons. SAMHD1 is a dNTPase that hydrolyzes dNTPs into deoxyribonucleosides (dNs) and triphosphates (48). It has been identified as a restriction factor that blocks infection by a broad range of retroviruses, including HIV-1, in noncycling myeloid-lineage cells and quiescent CD4+ T lymphocytes (49–54). Owing to its dNTPase activity, SAMHD1



can degrade the analog cytarabine triphosphate and reduce its concentrations in cells, such as the patient-derived acute myeloid leukemia blasts, thereby posing a significant barrier to the effective analog cytarabine-based treatment (45). However, SAMHD1 can hydrolyze several active triphosphate (TP) nucleoside analogs used for anti-cancer therapies (47). Therefore, evaluation of the expression levels of SAMHD1 in patients with stages II and III colorectal cancer before adjuvant chemotherapy is warranted.

SAMHD1 could complement MSI/MMR status as a promising molecular marker, leading to more accurate treatment decisions in patients with stage II colorectal cancer. The MSI/MMR status of the tumor is the only molecular marker involved in adjuvant chemotherapy decisions for stage II colorectal cancer (11). However, the MSI/MMR rate is 10%–15%, while 20% of patients with stage II colorectal cancer experience relapse after surgery (3).

This leads to the emerging need to identify novel biomarkers for the effective treatment of colorectal cancer. Our results show that SAMHD1 expression only partially overlaps with tumors defined by the MSI/MMR status. In this study, high expression of SAMHD1 was approximately 12% and conferred poor prognosis and less benefits from adjuvant chemotherapy for stage II disease in both the discovery and validation datasets. We will further assess the prognostic and predictive value of SAMHD1 using immunohistochemistry in a prospective multicenter cohort before clinical practice.

The major strength of our study is its nested case-control design combined with proteomics. The nested case-control design is an efficient method to identify novel prognostic biomarkers using the available, large sets of clinical data storing biological samples and taking both feasibility and economic factors into account (55). We



identified 21 proteins associated with the prognosis of patients with colorectal cancer by using quantitative proteomics in a nested case-control cohort within a large cohort of patients with colorectal cancer. Among these 21 proteins, five proteins showed a promising role as potential biomarkers for the identification of high-risk populations and chemo-sensitive patients with stage II colorectal cancer. Therefore, further studies are required to validate these results.

Each omics discipline has its own advantages and disadvantages, and can give information about many aspects of disease from transcriptomics signatures to proteomic profiles. By comparison, colorectal cancer-related protein-coding genes have little overlap with known cancer genes, this is one of the advantages of proteomics over other omics (29, 56). It is logical therefore to examine this extensive information in parallel with the aim of revealing those attributes that can be considered robust and sensitive enough to work as a biomarker of patient risk (57).

While the results are promising, this study has several limitations. Firstly, this study was lack of immunohistochemical validation of SAMHD1 due to the retrospective design, we could not obtain effective FFPE specimen from many patients because of the long storage time. We will further validate these results using the prospective, multicenter clinical trials. Secondly, since there were few stage III patients in our cohort, we do get lose the predictivity in stage III patients, the specific reasons are not clear, and further research is needed in the future. Thirdly, we did not perform more detailed analysis about adjuvant chemotherapy regimens due to the lack of specific treatment information in public datasets, so SAMHD1 should be a predictive parameter for a group of drugs, and much more clinical data should be available until SAMHD1 could be an add-on to clinical practice.

In conclusion, our research showed that SAMHD1 can effectively stratify patients with stage II colorectal cancer into subgroups with good and poor prognosis, thereby complementing the prognostic value of the MSI/MMR status that is used to evaluate the prognosis of these patients. Moreover, our results showed that the expression levels of SAMHD1 can identify stages II and III patients who could benefit from adjuvant chemotherapy. Thus, SAMHD1 may potentially be used as an easy and useful tool in clinical practice to develop more accurate treatment decisions for patients with stages II and III colorectal cancer.

## Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/[Supplementary material](#).

## Ethics statement

The studies involving human participants were reviewed and approved by Yunnan Cancer Hospital Ethics Committee (No. KY2019141). The patients/participants provided their written informed consent to participate in this study. Written informed

consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

## Author contributions

DY, SZ, and SY helped design the statistical approach, performed all the analyses, interpreted the results, and drafted the manuscript. YD, XC, LW, and WW helped with the collection of patient-related data. CL helped with the statistical methodology. TZ, ZL, and YH conceptualized and led the study, helped design the study and interpret the results, and revised and submitted the manuscript. All authors approved the final manuscript.

## Funding

This study was supported by the National Natural Scientific Foundation of China (82001986, 82073569, 81973147), National Science Fund for Distinguished Young Scholars (81925023), the Outstanding Youth Science Foundation of Yunnan Basic Research Project (202101AW070001, 202001AW070021), the Key Science Foundation of Yunnan Basic Research (202101AS070040), Yunnan Digitalization, Development and Application of Biotic Resource (202002AA100007), the Innovation Team of Kunming Medical University (CXTD202110).

## Acknowledgments

The authors thank the patients for making this study possible.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2022.939982/full#supplementary-material>

## References

- Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: Cancer J Clin* (2021) 71(3):209–49. doi: 10.3322/caac.21660
- Zhang S, Sun K, Zheng R, Zeng H, Wang S, Chen R, et al. Cancer incidence and mortality in China, 2015. *J Natl Cancer Center* (2021) 1(1):2–11. doi: 10.1016/j.jncc.2020.12.001
- Siegel RL, Miller KD, Goding Sauer A, Fedewa SA, Butterly LF, Anderson JC, et al. Colorectal cancer statistics, 2020. *CA: Cancer J Clin* (2020) 70(3):145–64. doi: 10.3322/caac.21601
- Manfredi S, Bouvier AM, Lepage C, Hatem C, Dancourt V, Faivre J. Incidence and patterns of recurrence after resection for cure of colonic cancer in a well defined population. *Br J Surg* (2006) 93(9):1115–22. doi: 10.1002/bjs.5349
- National-Comprehensive-Cancer-Network(NCCN) and Nccn Clinical Practice Guidelines in Oncology (Nccn Guidelines). *Rectal Cancer.(Version 1.2021)*. Fort Washington, PA: NCCN (2021).
- National-Comprehensive-Cancer-Network(NCCN) and Nccn Clinical Practice Guidelines in Oncology (Nccn Guidelines). *Colon Cancer.(Version 4.2021)*. Fort Washington, PA: NCCN (2021).
- Argiles G, Tabernero J, Labianca R, Hochhauser D, Salazar R, Iveson T, et al. Localised colon cancer: Esmo clinical practice guidelines for diagnosis, treatment and follow-up. *Ann Oncol* (2020) 31(10):1291–305. doi: 10.1016/j.annonc.2020.06.022
- Glynn-Jones R, Wyrwicz L, Tiret E, Brown G, Rodel C, Cervantes A, et al. Rectal cancer: Esmo clinical practice guidelines for diagnosis, treatment and follow-up. *Ann Oncol* (2017) 28:iv22–iv40. doi: 10.1093/annonc/mdx224
- Andre T, de Gramont A, Vernerey D, Chibaudel B, Bonnetain F, Tijeras-Raballand A, et al. Adjuvant fluorouracil, leucovorin, and oxaliplatin in stage ii to iii colon cancer: Updated 10-year survival and outcomes according to braf mutation and mismatch repair status of the mosaic study. *J Clin Oncol* (2015) 33(35):4176–87. doi: 10.1200/JCO.2015.63.4238
- Figueredo A, Coombes ME, Mukherjee S. Adjuvant therapy for completely resected stage ii colon cancer. *Cochrane Database Syst Rev* (2008) 3:CD005390. doi: 10.1002/14651858.CD005390.pub2
- Parent P, Cohen R, Rassy E, Svrcek M, Taieb J, Andre T, et al. A comprehensive overview of promising biomarkers in stage ii colorectal cancer. *Cancer Treat Rev* (2020) 88:102059. doi: 10.1016/j.ctrv.2020.102059
- Baxter NN, Kennedy EB, Bergsland E, Berlin J, George TJ, Gill S, et al. Adjuvant therapy for stage ii colon cancer: Asco guideline update. *J Clin Oncol* (2021) 40(8):892–910. doi: 10.1200/JCO.21.02538
- Webber EM, Lin JS, Evelyn PW. Oncotype dx tumor gene expression profiling in stage ii colon cancer. application: Prognostic, risk prediction. *PLoS Curr* (2010) 2:RRN1177. doi: 10.1371/currents.RRN1177
- Niedzwiecki D, Frankel WL, Venook AP, Ye X, Friedman PN, Goldberg RM, et al. Association between results of a gene expression signature assay and recurrence-free interval in patients with stage ii colon cancer in cancer and leukemia group b 9581 (Alliance). *J Clin Oncol* (2016) 34(25):3047–53. doi: 10.1200/JCO.2015.65.4699
- Kopetz S, Tabernero J, Rosenberg R, Jiang ZQ, Moreno V, Bachleitner-Hofmann T, et al. Genomic classifier coloproct predicts recurrence in stage ii colorectal cancer patients more accurately than clinical factors. *Oncologist* (2015) 20(2):127–33. doi: 10.1634/theoncologist.2014-0325
- Zhang JX, Song W, Chen ZH, Wei JH, Liao YJ, Lei J, et al. Prognostic and predictive value of a microRNA signature in stage ii colon cancer: A microRNA expression analysis. *Lancet Oncol* (2013) 14(13):1295–306. doi: 10.1016/S1470-2045(13)70491-1
- Tie J, Wang Y, Tomasetti C, Li L, Springer S, Kinde I, et al. Circulating tumor DNA analysis detects minimal residual disease and predicts recurrence in patients with stage ii colon cancer. *Sci Transl Med* (2016) 8(346):346ra92. doi: 10.1126/scitranslmed.aaf6219
- Reinert T, Henriksen TV, Christensen E, Sharma S, Salari R, Sethi H, et al. Analysis of plasma cell-free DNA by ultradeep sequencing in patients with stages I to iii colorectal cancer. *JAMA Oncol* (2019) 5(8):1124–31. doi: 10.1001/jamaoncol.2019.0528
- Tie J, Cohen JD, Wang Y, Christie M, Simons K, Lee M, et al. Circulating tumor DNA analyses as markers of recurrence risk and benefit of adjuvant therapy for stage iii colon cancer. *JAMA Oncol* (2019) 5(12):1710–7. doi: 10.1001/jamaoncol.2019.3616
- Cai G, Cai M, Feng Z, Liu R, Liang L, Zhou P, et al. A multilocus blood-based assay targeting circulating tumor DNA methylation enables early detection and early relapse prediction of colorectal cancer. *Gastroenterology* (2021) 161(6):2053–6.e2. doi: 10.1053/j.gastro.2021.08.054
- Chen G, Peng J, Xiao Q, Wu HX, Wu X, Wang F, et al. Postoperative circulating tumor DNA as markers of recurrence risk in stages ii to iii colorectal cancer. *J Hematol Oncol* (2021) 14(1):80. doi: 10.1186/s13045-021-01089-z
- Pagès F, Mlecnik B, Marliot F, Bindea G, Ou F-S, Bifulco C, et al. International validation of the consensus immunoscore for the classification of colon cancer: A prognostic and accuracy study. *Lancet* (2018) 391(10135):2128–39. doi: 10.1016/s0140-6736(18)30789-x
- Reichling C, Taieb J, Derangere V, Klopffenstein Q, Le Malicot K, Gornet JM, et al. Artificial intelligence-guided tissue analysis combined with immune infiltrate assessment predicts stage iii colon cancer outcomes in Petacc08 study. *Gut* (2020) 69(4):681–90. doi: 10.1136/gutjnl-2019-319292
- Sanz-Pamplona R, Melas M, Maoz A, Schmit SL, Rennert H, Lejbkovicz F, et al. Lymphocytic infiltration in stage ii microsatellite stable colorectal tumors: A retrospective prognosis biomarker analysis. *PLoS Med* (2020) 17(9):e1003292. doi: 10.1371/journal.pmed.1003292
- Skrede O-J, De Raedt S, Kleppe A, Hveem TS, Liestøl K, Maddison J, et al. Deep learning for prediction of colorectal cancer outcome: A discovery and validation study. *Lancet* (2020) 395(10221):350–60. doi: 10.1016/s0140-6736(19)32998-8
- Lee JM, Kohn EC. Proteomics as a guiding tool for more effective personalized therapy. *Ann Oncol* (2010) 21 Suppl 7:vii205–10. doi: 10.1093/annonc/mdq375
- Torres S, Garcia-Palmero I, Herrera M, Bartolome RA, Pena C, Fernandez-Acenero MJ, et al. Loxl2 is highly expressed in cancer-associated fibroblasts and associates to poor colon cancer survival. *Clin Cancer Res* (2015) 21(21):4892–902. doi: 10.1158/1078-0432.CCR-14-3096
- Jiang Y, Sun A, Zhao Y, Ying W, Sun H, Yang X, et al. Proteomics identifies new therapeutic targets of early-stage hepatocellular carcinoma. *Nature* (2019) 567(7747):257–61. doi: 10.1038/s41586-019-0987-8
- Vasaikar S, Huang C, Wang X, Petyuk VA, Savage SR, Wen B, et al. Proteogenomic analysis of human colon cancer reveals new therapeutic opportunities. *Cell* (2019) 177(4):1035–49 e19. doi: 10.1016/j.cell.2019.03.030
- Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika* (1983) 70(1):41–55. doi: 10.1093/biomet/70.1.41
- D'Agostino RB Jr. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Stat Med* (1998) 17(19):2265–81. doi: 10.1002/(sici)1097-0258(19981015)17:19<2265::aid-sim918>3.0.co;2-b
- Wisniewski JR, Zougman A, Nagaraj N, Mann M. Universal sample preparation method for proteome analysis. *Nat Methods* (2009) 6(5):359–62. doi: 10.1038/nmeth.1322
- Kaur P, Porras TB, Ring A, Carpten JD, Lang JE. Comparison of tcga and genie genomic datasets for the detection of clinically actionable alterations in breast cancer. *Sci Rep* (2019) 9(1):1482. doi: 10.1038/s41598-018-37574-8
- Cheadle C, Vawter MP, Freed WJ, Becker KG. Analysis of microarray data using z score transformation. *J Mol Diagnostics* (2003) 5(2):73–81. doi: 10.1016/s1525-1578(10)60455-2
- Camp RL, Dolled-Filhart M, Rimm DL. X-Tile: A new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization. *Clin Cancer Res* (2004) 10(21):7252–9. doi: 10.1158/1078-0432.ccr-04-0713
- Knecht KM, Buzovetsky O, Schneider C, Thomas D, Srikanth V, Kaderali L, et al. The structural basis for cancer drug interactions with the catalytic and allosteric sites of Samhd1. *Proc Natl Acad Sci U.S.A.* (2018) 115(43):E10022–E31. doi: 10.1073/pnas.1805593115
- Rassidakis GZ, Herold N, Myrberg IH, Tsesmetzis N, Rudd SG, Henter JI, et al. Low-level expression of Samhd1 in acute myeloid leukemia (Aml) blasts correlates with improved outcome upon consolidation chemotherapy with high-dose cytarabine-based regimens. *Blood Cancer J* (2018) 8(11):98. doi: 10.1038/s41408-018-0134-z
- Oellerich T, Schneider C, Thomas D, Knecht KM, Buzovetsky O, Kaderali L, et al. Selective inactivation of hypomethylating agents by Samhd1 provides a rationale for therapeutic stratification in aml. *Nat Commun* (2019) 10(1):3475. doi: 10.1038/s41467-019-11413-4
- Nadeu F, Martin-Garcia D, Clot G, Diaz-Navarro A, Duran-Ferrer M, Navarro A, et al. Genomic and epigenomic insights into the origin, pathogenesis, and clinical behavior of mantle cell lymphoma subtypes. *Blood* (2020) 136(12):1419–32. doi: 10.1182/blood.202005289

40. Rentoft M, Lindell K, Tran P, Chabes AL, Buckland RJ, Watt DL, et al. Heterozygous colon cancer-associated mutations of Samhd1 have functional significance. *Proc Natl Acad Sci USA* (2016) 113(17):4723–8. doi: 10.1073/pnas.1519128113
41. Yang CA, Huang HY, Chang YS, Lin CL, Lai IL, Chang JG. DNA-Sensing and nuclease gene expressions as markers for colorectal cancer progression. *Oncology* (2017) 92(2):115–24. doi: 10.1159/000452281
42. Chen Z, Hu J, Ying S, Xu A. Dual roles of Samhd1 in tumor development and chemoresistance to anticancer drugs. *Oncol Lett* (2021) 21(6):451. doi: 10.3892/ol.2021.12712
43. Xagoraris I, Vassilakopoulos TP, Drakos E, Angelopoulou MK, Panitsas F, Herold N, et al. Expression of the novel tumour suppressor sterile alpha motif and hd domain-containing protein 1 is an independent adverse prognostic factor in classical Hodgkin lymphoma. *Br J Haematol* (2021) 193(3):488–96. doi: 10.1111/bjh.17352
44. Uhlen M, Zhang C, Lee S, Sjöstedt E, Fagerberg L, Bidkhorji G, et al. A pathology atlas of the human cancer transcriptome. *Science* (2017) 357(6352): ean2507. doi: 10.1126/science.aan2507
45. Schneider C, Oellerich T, Baldauf HM, Schwarz SM, Thomas D, Flick R, et al. Samhd1 is a biomarker for cytarabine response and a therapeutic target in acute myeloid leukemia. *Nat Med* (2017) 23(2):250–5. doi: 10.1038/nm.4255
46. Clifford R, Louis T, Robbe P, Ackroyd S, Burns A, Timbs AT, et al. Samhd1 is mutated recurrently in chronic lymphocytic leukemia and is involved in response to DNA damage. *Blood* (2014) 123(7):1021–31. doi: 10.1182/blood-2013-04-490847
47. Ordóñez P, Kunzelmann S, Groom HC, Yap MW, Weising S, Meier C, et al. Samhd1 enhances nucleoside-analogue efficacy against hiv-1 in myeloid cells. *Sci Rep* (2017) 7:42824. doi: 10.1038/srep42824
48. Goldstone DC, Ennis-Adeniran V, Hedden JJ, Groom HC, Rice GI, Christodoulou E, et al. Hiv-1 restriction factor Samhd1 is a deoxynucleoside triphosphate triphosphohydrolase. *Nature* (2011) 480(7377):379–82. doi: 10.1038/nature10623
49. White TE, Brandariz-Núñez A, Valle-Casuso JC, Amie S, Nguyen LA, Kim B, et al. The retroviral restriction ability of Samhd1, but not its deoxynucleotide triphosphohydrolase activity, is regulated by phosphorylation. *Cell Host Microbe* (2013) 13(4):441–51. doi: 10.1016/j.chom.2013.03.005
50. Descours B, Cribier A, Chable-Bessia C, Ayinde D, Rice G, Crow Y, et al. Samhd1 restricts hiv-1 reverse transcription in quiescent Cd4(+) T-cells. *Retrovirology* (2012) 9:87. doi: 10.1186/1742-4690-9-87
51. Baldauf HM, Pan X, Erikson E, Schmidt S, Daddacha W, Burggraf M, et al. Samhd1 restricts hiv-1 infection in resting Cd4(+) T cells. *Nat Med* (2012) 18(11):1682–7. doi: 10.1038/nm.2964
52. Laguette N, Sobhian B, Casartelli N, Ringear M, Chable-Bessia C, Ségéral E, et al. Samhd1 is the dendritic- and myeloid-Cell-Specific hiv-1 restriction factor counteracted by vpx. *Nature* (2011) 474(7353):654–7. doi: 10.1038/nature10117
53. Hrecka K, Hao C, Gierszewska M, Swanson SK, Kesik-Brodacka M, Srivastava S, et al. Vpx relieves inhibition of hiv-1 infection of macrophages mediated by the Samhd1 protein. *Nature* (2011) 474(7353):658–61. doi: 10.1038/nature10195
54. Berger A, Sommer AF, Zwarg J, Hamdorf M, Welzel K, Esly N, et al. Samhd1-deficient Cd14+ cells from individuals with aicardi-goutières syndrome are highly susceptible to hiv-1 infection. *PLoS Pathog* (2011) 7(12):e1002425. doi: 10.1371/journal.ppat.1002425
55. Graziano F, Valsecchi MG, Rebora P. Sampling strategies to evaluate the prognostic value of a new biomarker on a time-to-Event end-point. *BMC Med Res Methodol* (2021) 21(1):93. doi: 10.1186/s12874-021-01283-0
56. Imperial R, Ahmed Z, Toor OM, Erdoğan C, Khaliq A, Case P, et al. Comparative proteogenomic analysis of right-sided colon cancer, left-sided colon cancer and rectal cancer reveals distinct mutational profiles. *Mol Cancer* (2018) 17(1):177. doi: 10.1186/s12943-018-0923-9
57. Kane LE, Mellotte GS, Conlon KC, Ryan BM, Maher SG. Multi-omic biomarkers as potential tools for the characterisation of pancreatic cystic lesions and cancer: Innovative patient data integration. *Cancers (Basel)* (2021) 13(4):769. doi: 10.3390/cancers13040769