

2023

## CSRP1 expression is associated with a mesenchymal, stroma-rich tumor profile and poor prognosis in colon cancer

SEÇİL DEMİRKOL CANLI

Follow this and additional works at: <https://journals.tubitak.gov.tr/medical>



Part of the [Medical Sciences Commons](#)

### Recommended Citation

CANLI, SEÇİL DEMİRKOL (2023) "CSRP1 expression is associated with a mesenchymal, stroma-rich tumor profile and poor prognosis in colon cancer," *Turkish Journal of Medical Sciences*: Vol. 53: No. 6, Article 16. <https://doi.org/10.55730/1300-0144.5736>

Available at: <https://journals.tubitak.gov.tr/medical/vol53/iss6/16>

This Article is brought to you for free and open access by TÜBİTAK Academic Journals. It has been accepted for inclusion in Turkish Journal of Medical Sciences by an authorized editor of TÜBİTAK Academic Journals. For more information, please contact [academic.publications@tubitak.gov.tr](mailto:academic.publications@tubitak.gov.tr).

## CSR1 expression is associated with a mesenchymal, stroma-rich tumor profile and poor prognosis in colon cancer

Seçil DEMİRKOL CANLI 

Division of Tumor Pathology, Department of Clinical Oncology, Cancer Institute, Hacettepe University, Ankara, Türkiye

Received: 20.08.2023 • Accepted/Published Online: 31.10.2023 • Final Version: 12.12.2023

**Background/aim:** Cysteine and glycine-rich protein 1 (CSR1) is involved in the cysteine-rich protein family and is a marker of smooth muscle lineages. In colon cancer, the expression of this gene is associated with poor prognosis. In this study, the aim was to reevaluate its prognostic relationship in independent cohorts and explore potential underlying biological processes that are linked to aggressive behavior in tumors with high CSR1 expression, such as epithelial-to-mesenchymal transition (EMT), stromal fractions in the tumor microenvironment, and consensus molecular subtypes (CMSs).

**Materials and methods:** RNA sequencing (RNAseq)-, microarray-, and single-cell RNAseq (scRNAseq)-based transcriptomic data were obtained from public databases. The EMT score was calculated based on the expression of E-cadherin and vimentin genes using a previously published method. The stromal score generated by the ESTIMATE method was utilized for the analysis of correlation with the CSR1 expression. The scRNAseq data were analyzed via the Seurat R package. The immunohistochemistry-based protein level expression of CSR1 was evaluated using the Human Protein Atlas database.

**Results:** Lower CSR1 expression was noted in colon tumors compared to normal colon tissue. Patients with a high CSR1 expression had shorter recurrence-free, overall, and disease-specific survivals in the GSE39582 and GSE17536 datasets ( $p < 0.05$ ). The methylation level of the CSR1 gene was negatively correlated ( $r = -0.57$ ,  $p < 0.0001$ ) with CSR1 expression in The Cancer Genome Atlas colon adenocarcinoma dataset. CSR1 expression was positively correlated with the expression of mesenchymal markers, EMT score, and stromal score obtained via the ESTIMATE method. CMS4 colon tumors had a significantly higher CSR1 expression compared to other CMSs. Analysis of the scRNAseq data revealed that CSR1 was expressed by epithelial cells and cancer-associated fibroblasts in the colorectal tumor microenvironment, which was also confirmed by the protein expression data from the Human Protein Atlas database.

**Conclusion:** CSR1 expression is associated with CMS4, a more mesenchymal stroma-rich molecular profile, and poor prognosis in colon cancer.

**Key words:** Colon cancer, bioinformatics, gene expression, CSR1, prognosis

### 1. Introduction

Cysteine and glycine-rich protein 1 (CSR1), a member of the cysteine-rich protein family, is a LIM domain-containing protein involved in smooth muscle differentiation [1]. It is considered as a well-known marker for smooth muscle lineages [2]. CSR1 has a role in actin filament bundling by cross-linking actin filaments and stabilizing the interaction of alpha-actinin with actin filament bundles [3]. In healthy tissues, the highest transcript level of CSR1 was reported in the prostate and colon, followed by the brain and testis [4]. In human skin fibroblasts, CSR1 expression was induced by growth-inhibitory signals, and was associated with a differentiated morphology towards the myofibroblast lineage [5]. The ectopic expression of CSR1 in these cells was reported to suppress cell proliferation [5]. Supporting these findings,

another study showed that transforming growth factor-beta (TGF- $\beta$ 1) increased the expression of CSR1 at the protein level concomitant with an increase in myofibroblast differentiation [6]. Higher CSR1 expression was also reported in lung fibrosis compared to control lung tissue [6].

However, relatively little is known about its function and potential implications in cancer. In hepatocellular carcinoma (HCC), 56% of cases had aberrant methylation and a lower expression of CSR1 compared to normal liver tissue [7]. CSR1 was listed among hub genes that are downregulated in prostate cancer compared to benign prostate hyperplasia [8]. Moreover, high CSR1 expression was associated with longer disease-free survival (DFS) in prostate cancer [8]. In colon cancer, CSR1 has been listed in a 3-gene list generated to distinguish colon cancer from

\* Correspondence: [secil.demirkol@hacettepe.edu.tr](mailto:secil.demirkol@hacettepe.edu.tr)

normal colon tissue with 90.32% accuracy [9]. CSRPI expression was downregulated in 15 of 19 colorectal tumors; thus, it was suggested as a potential tumor suppressor in colorectal cancer (CRC) [10]. A recent study showed that a high expression of CSRPI was associated with poor prognosis in The Cancer Genome Atlas (TCGA) colon adenocarcinoma (COAD) cohort [11]. However, evaluation of this gene as a prognostic marker has not yet been performed in independent cohorts.

In this study, the aim was to assess the biological and clinical parameters that are associated with CSRPI expression in colon cancer. For this purpose, publicly available transcriptomic data from colon tumors and normal colon tissues were analyzed in parallel with previously generated molecular scoring and subgrouping methods that reflected the epithelial to mesenchymal transition (EMT) status, stromal fractions, and cancer-associated fibroblast (CAF) involvement. The findings herein showed that the CSRPI expression was downregulated in colon cancer compared to normal colon tissue. Furthermore, the CSRPI expression and methylation levels were negatively correlated in colon tumors. A high CSRPI expression was associated with a mesenchymal and stroma-rich molecular profile and poor prognosis. Single-cell RNA sequencing (scRNAseq) and immunohistochemistry (IHC) data indicated that CSRPI was expressed by epithelial and stromal cells in the tumor microenvironment.

## 2. Materials and methods

### 2.1. In silico analysis of the gene expression and methylation levels

Raw microarray-based transcriptomic data for primary COAD, GSE39582 (n = 566) and GSE17536 (n = 177) were downloaded from the Gene Expression Omnibus (GEO) database (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi>) and RMA normalized [12,13]. An EMT score for samples in the GSE39582 and GSE17536 datasets were generated utilizing a previously published formula based on the expression of E-cadherin and vimentin (VIM) genes [14]. The EMT score has a range between -2 and 0 from the most epithelial to the most mesenchymal phenotype. Consensus molecular subtype (CMS) information was obtained from the Synapse web portal ([www.synapse.org](http://www.synapse.org)). Samples that could not be classified in any of the 4 CMSs were not used for comparison of the CSRPI expression across the CMSs. A previously defined method, ESTIMATE, was utilized to obtain stromal scores inferring the fraction of stromal cells in the tumor samples [15]. ESTIMATE stromal scores for the GSE39582 dataset were generated using the ESTIMATE R package (University of Texas MD Anderson Cancer Center; <https://bioinformatics.mdanderson.org/estimate/>

[rpackage.html](https://bioinformatics.mdanderson.org/estimate/disease.html)). ESTIMATE stromal scores based on the TCGA COAD RNAseq V2 data were downloaded (<https://bioinformatics.mdanderson.org/estimate/disease.html>). To define the tumor subgroups based on the CAF marker expression in the GSE39582 and GSE17536 datasets, a published list was used [16].

TCGA COAD primary tumor RNAseq data were downloaded from the National Cancer Institute Genomic Data Commons (GDC) portal (<https://portal.gdc.cancer.gov/>) in STAR-counts format. Formalin-fixed paraffin-embedded (FFPE) tissues were filtered out. For duplicate samples, the sample with the smaller total read count was excluded from the study. The remaining 451 unique tumors were used in the following steps. Genes with a count value of zero in more than 90% of the samples were removed from the count matrix. Counts were normalized using the DESeq2 package and log transformed [17]. 186 tumors had both the RNAseq-based gene expression data and ESTIMATE scores, which were used for the correlation analysis of the CSRPI gene expression and stromal score in TCGA.

The linear correlation of CSRPI expression and methylation status was studied via cBioPortal for Cancer Genomics (<https://www.cbioportal.org/>) [18,19]. Infinium Human Methylation 450K BeadChip and RNA expression (RNAseq v2) data in Colorectal Adenocarcinoma (TCGA, Firehose Legacy) datasets were used. Methylation beta and log(RSEM+1) values were utilized for the analysis of linear correlation. A comparison of the expression of CSRPI in colorectal tumors and normal colon and rectum tissues was conducted via the Gene Expression Profiling Interactive Analysis (GEPIA) platform (<https://gepia.cancer-pku.cn/>) [20].

### 2.2. Log-rank-based prognostic analysis at multiple cut-off values

Cut-off-based evaluation of the prognostic relationships for the CSRPI expression were conducted based on log-rank multiple cut-off graphs, as described previously [21]. Briefly, the patients were divided into 2 groups with a high and low expression of CSRPI using each value of the CSRPI expression as the cut-off value. Log-rank tests were performed based on the groups defined by each cut-off value. Log-rank p-values (y-axis) were plotted against all of the cut-off values (x-axis). Red and blue indicate the hazard ratio (HR) above and below 1, respectively, using the low expression group as a reference. The cut-off values with the lowest log-rank p-value within the interquartile range were used to define the groups with high and low expression for the Kaplan–Meier graphs.

### 2.3. scRNAseq data analysis

scRNAseq data of colorectal primary tumors were obtained from the GSE178318 dataset. Three patients who received no treatment were included in the analysis. An

aligned raw gene expression matrix was downloaded from the GEO database and processed via the Seurat package in R Bioconductor [22]. Quality filtering was performed as follows: cells with less than 500 or more than 6000 expressed features, and cells with mitochondrial gene expression greater than 15% were filtered out. Doublets were detected and removed using the DoubletFinder package [23]. Next, 2000 variable features were determined using the FindVariableFeatures function. The uniform manifold approximation and projection method was used for dimension reduction. The list of cell-specific markers that was presented by Che et al. [24] was used to define the cellular clusters, with minor modifications. The NKG7 gene was added to the list of natural killer (NK) cell markers. The cell cluster that expressed CD68, CD163, CD14, and LYZ was named Myeloid cells.

#### 2.4. Evaluation of the IHC-based CSRPI expression

IHC data available in the Human Protein Atlas were used [25, 26]. IHC-based staining for CSRPI (clone: HPA045617) was evaluated in colon tumors. The consensus dataset of the RNA expression obtained from healthy tissues was utilized for the assessment of the CSRPI expression.

#### 2.5. Statistical analysis

Kaplan–Meier graphs were generated using Graphpad Prism 8 (San Diego, CA, USA). One-way analysis of variance (ANOVA), Tukey's multiple comparison tests, unpaired t tests, and Pearson correlation analysis were performed using Graphpad Prism 8. Cox univariate and multivariate regression analysis (MVA) were performed using IBM SPSS Statistics for Windows 23.0 (IBM Corp., Armonk, NY, USA). Patients with nonzero survival data were included in these analyses. Documented recurrence (0 = no, 1 = yes) was used as status for DFS.

### 3. Results

#### 3.1. CSRPI expression in malignant and benign tissues

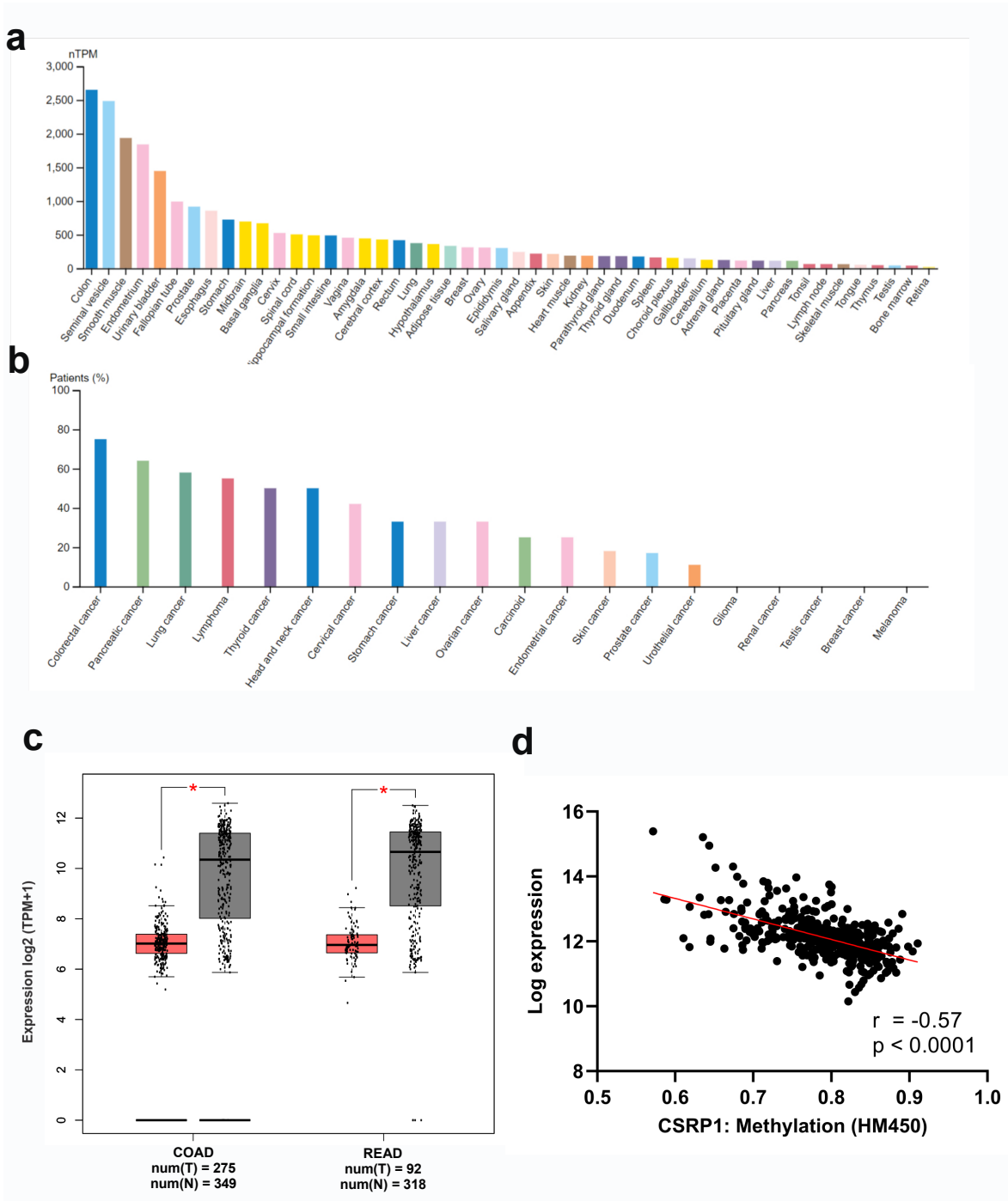
The CSRPI gene expression was evaluated in various cancer types and healthy tissues using publicly available datasets. Based on the consensus RNAseq dataset of the Human Protein Atlas, the expression of CSRPI was the highest in colon tissue among the 54 healthy tissues available (Figure 1a). At the protein level, measured via IHC, the percentage of patients with moderate or high expression was the highest in CRC among the 20 cancer types (Figure 1b). When the expression of CSRPI was analyzed via the GEPIA web tool, a significantly lower expression of CSRPI was noted in both colon and rectal tumors (TCGA) compared to normal colon and rectum tissue (TCGA+GTEx) (log fold change >3,  $p < 0.05$ ) (Figure 1c). These data showed that the CSRPI gene was expressed at high levels in both the malignant and normal colonic tissues, and its expression was lower in colon tumors compared to normal colon tissue.

DNA methylation is one of the main epigenetic mechanisms of regulation of gene expression levels and CSRPI was shown to be inactivated via aberrant methylation in HCC [7]. To understand whether DNA methylation could be a potential mechanism that is involved in the regulation of CSRPI expression in CRC, publicly available CSRPI methylation and expression data from TCGA colorectal adenocarcinoma cohort were analyzed. Indeed, the methylation and expression levels were strongly negatively correlated ( $r = -0.57$ ,  $p < 0.0001$ ) for CSRPI, suggesting that DNA methylation may have a role in the transcriptional regulation of this gene (Figure 1d).

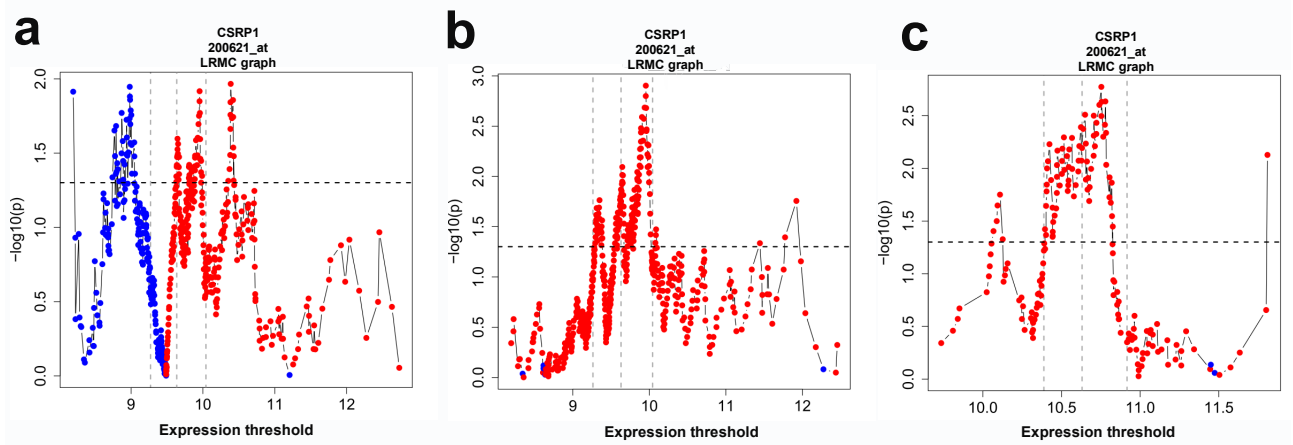
#### 3.2. Association of the CSRPI expression with prognosis and clinicopathological parameters

A recent study identified that a high CSRPI expression was an independent marker of poor prognosis (HR=1.895, 95% CI: 1.078–3.330,  $p = 0.026$ ) in TCGA COAD dataset [11]. In order to evaluate the robustness of CSRPI as a prognostic marker across multiple cohorts, independent transcriptomic datasets were analyzed. A high expression of CSRPI was significantly associated with poor prognosis, consistently, at multiple cut-off values within the interquartile range of the CSRPI expression in the GSE39582 (Figures 2a and 2b) and GSE17536 (Figure 2c) datasets. A high CSRPI expression was associated significantly ( $p < 0.05$ ) with overall survival (OS) (HR = 1.442, 95% CI: 1.054–1.973) (Figure 3a) and recurrence-free survival (RFS) (HR = 1.691, 95% CI: 1.181–2.422) in the GSE39582 dataset (Figure 3b), and disease-specific survival (DSS) (HR = 2.295, 95% CI: 1.331–3.959) in the GSE17536 dataset (Figure 3c). No significant association was noted with disease free survival in the GSE17536 dataset. Overall, a high CSRPI expression was associated with poor prognosis in 2 independent cohorts when tested with 3 different measures of clinical outcome (OS, RFS, and DSS).

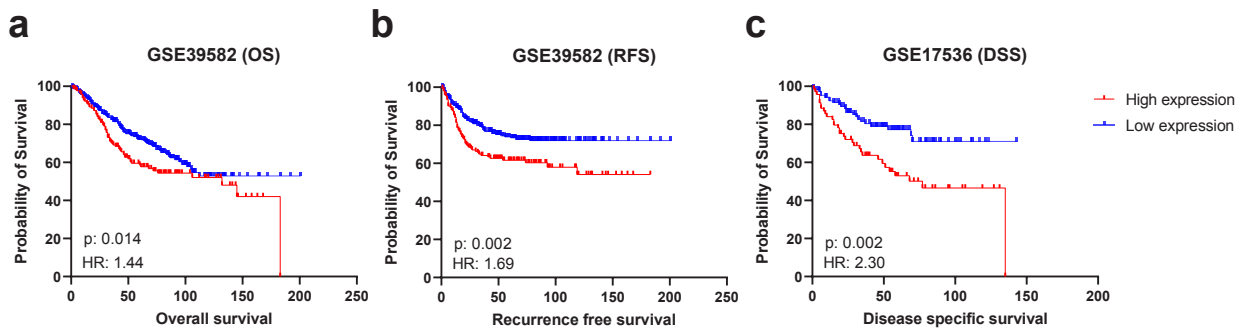
To assess whether CSRPI is an independent prognostic marker, MVA was performed with the confounding factors identified previously in the GSE39582 dataset [21]. The results revealed that the CSRPI expression was a significant predictor of poor prognosis independent of the microsatellite instability (MSI) status, TNM stage, and KRAS-BRAF mutation status (Table 1). Although the differences between the mean expression levels were minor (the mean expression was 9.35, 9.72, 9.80, and 9.69 for stages 1, 2, 3, and 4, respectively), the CSRPI expression was significantly lower in stage 1 compared to stages 2 and 3 (Figure 4a). When the CSRPI expression was compared across tumors with varying clinical characteristics, there was no significant difference between tumors with different grades (Figure 4b), the MSI status (Figure 4c), and tumors with and without the KRAS or BRAF mutation (Figure 4d).



**Figure 1.** Expression and methylation of CSR1 in the healthy and malignant tissues. RNAseq-based expression of CSR1 in healthy tissues within the consensus dataset of Human Protein Atlas (a) and the percentage of patients with a moderate or high protein level of CSR1 expression across the available cancer types (b) are shown. The expression of CSR1 in colon and rectal tumors (TCGA) compared to a normal colon and rectum (TCGA+GTEx) were plotted using the GEPIA web server (c). \*  $p < 0.05$ . Correlation of CSR1 expression and methylation data obtained from cBioportal database ( $n = 372$ ) (d). Pearson  $r$ - and  $p$ -values are indicated.



**Figure 2.** Log-rank multiple cut-off graphs for the CSR1 expression. Log-rank test results are plotted at all cut-off values: the graphic indicates the log-rank values (shown as dots) obtained at every possible cut-off for CSR1 in GSE39582 (a-b) and GSE17536 (c) for OS (a), RFS (b) and DSS (c). Red and blue indicate that high CSR1 expression is associated with poor (HR > 1) and good prognosis (HR < 1), respectively. Horizontal dotted line  $p = 0.05$ . Vertical dotted lines: (from left to right) first 25th percentile, median, and 75th percentile.



**Figure 3.** Kaplan-Meier plots for the CSR1 expression in GSE39582 and GSE17536. Prognostic comparison in GSE39582 and GSE17536 are shown in (a), (b), and (c), respectively. OS (a) and RFS (b) were used as clinical end-points for GSE39582. DSS was used for GSE17536 (c). Hazard ratios and log-rank p-values are indicated.

### 3.3. Association of the CSR1 expression with a mesenchymal, stroma rich molecular profile

Previous studies have indicated a strong association between EMT and the ability of cancer cells to invade other tissues and it is a key mechanism contributing to the progression of CRC [27]. As a high CSR1 expression was associated with a shorter RFS, next, the relationships between the expression of CSR1 and mesenchymal markers (SNAI1, SNAI2, ZEB1, ZEB2, TWIST1, TWIST2, and VIM) were studied. The CSR1 expression had a significant positive correlation with all 7 of the mesenchymal markers tested ( $p < 0.001$ ) in both the GSE39582 and TCGA COAD datasets (Table 2). The mean of the 2 r-values from these datasets were above 0.40 for 4 of the genes, ZEB1, ZEB2, TWIST2, and VIM (Table 2), indicating a strong correlation of the CSR1

expression with the mesenchymal markers. These findings were confirmed by applying a previously published EMT scoring method (see the Materials and methods section) in the GSE39582 and GSE17536 datasets. The CSR1 expression showed a significant positive correlation with the EMT score ( $r > 0.50$ ,  $p < 0.0001$ ) (Figures 5a and 5b). Overall, these data revealed that the CSR1 expression was correlated with the mesenchymal gene expression, suggesting an association with relatively mesenchymal subtypes of colon cancer.

The Colorectal Cancer Subtyping Consortium defined CMS4-type tumors as a mesenchymal tumor subtype that harbors prominent transforming growth factor  $\beta$  activation, and angiogenesis [28]. Therefore, the CSR1 expression was next evaluated among the CMSs. The CMS4-type tumors had a significantly higher CSR1

**Table 1.** Multivariate analysis in GSE39582 (RFS).

	No. of patients	Percentage	HR	95% CI for HR		p-value
				Lower	Upper	
MSI status						
pMMR (ref.)	444	78.45%	0.43	0.23	0.82	<b>0.011</b>
dMMR	75	13.30%				
NA	47	8.30%				
TNM stage*						
0	4	0.70%	1.66	1.27	2.16	<b>&lt;0.001</b>
1	33	5.83%				
2	264	46.64%				
3	205	36.20%				
4	60	10.60%				
KRAS or BRAF mutation						
Mutant**	268	47.35%	1.37	0.95	1.99	0.095
Wild type (ref.)	255	45.05%				
NA	43	7.60%				
CSRPI expression						
High	170	30.04%	1.94	1.35	2.80	<b>&lt;0.001</b>
Low (ref.)	396	69.96%				

\* Treated as a continuous variable (1, 2, 3, and 4), \*\* mutant for KRAS or BRAF. Reference groups are indicated (ref.) for the categorical variables. pMMR: proficient mismatch repair, dMMR: deficient mismatch repair, HR: hazard ratio, CI: confidence interval.

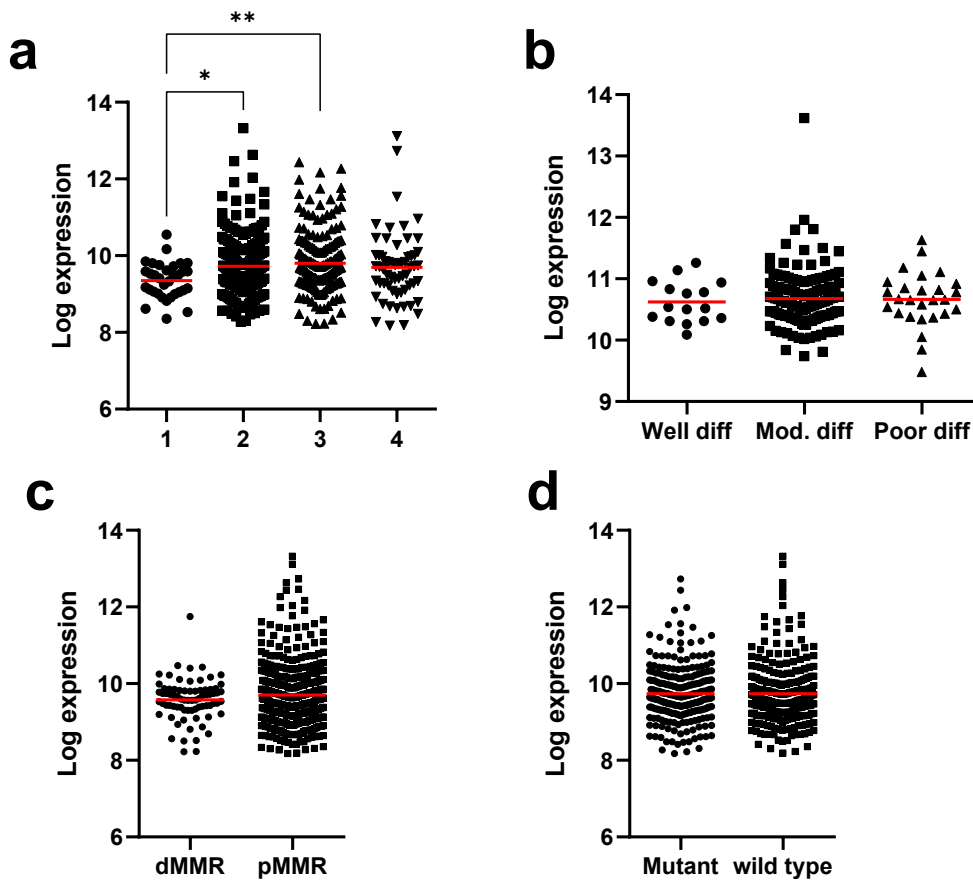
expression when compared to the other 3 types, CMS1, CMS2, and CMS3, in both the GSE39582 (Figure 5c) and TCGA COAD datasets (Figure 5d). The CMS4-type was also associated with a higher stromal invasion [28]. Accordingly, a previously developed method, ESTIMATE, was utilized to assess whether the stromal fractions in the tumor microenvironment are correlated with the CSRPI expression. The stromal scores generated by ESTIMATE, inferring the presence of stroma in tumor tissue, were significantly correlated with the CSRPI expression in both the GSE39582 (Figure 5e) and TCGA COAD datasets ( $r > 0.45$ ,  $p < 0.0001$ ) (Figure 5f). Taken together, these findings indicated that tumors with a high CSRPI expression had a stroma-rich profile that overlapped, to a large extent, with a CMS4 phenotype.

### 3.4. Epithelial cells and CAFs express CSRPI in the colon tumor microenvironment

In order to identify the cell types that express CSRPI, a scRNAseq dataset (GSE178318) of colorectal tumors was analyzed. The CSRPI expression was detected mainly in epithelial cells and CAFs, but very low or no expression was noted in cell clusters that included T cells, myeloid cells, B cells, NK cells, plasma, and mast cells (Figures 6a–6c). Among all of the cell types, CAFs expressed CSRPI at the

highest level, which is in line with the higher expression of CSRPI in tumors with higher stromal fractions. Furthermore, the CSRPI expression was analyzed within previously defined CRC subgroups based on the expression of CAF markers [16]. The analysis showed that tumors with a higher CAF marker expression (high group) had increased CSRPI expression, which gradually decreased in tumors with intermediate and low CAF marker expression (Figure 6d) ( $p < 0.0001$ ).

Evaluation of the IHC-based expression of CSRPI in colon tumors available in the Human Protein Atlas showed that 4, 5, and 2 cases had high, intermediate, and low expression of CSRPI, respectively (<https://www.proteinatlas.org/ENSG00000159176-CSRPI/pathology/colorectal+cancer>). No expression was detected for 1 patient. These data indicate that CSRPI was, indeed, expressed at the protein level in the colon tumors, and the level of expression varied among the cases. The subcellular localization was in cytoplasmic/membranous regions for all of the patients. The CSRPI expression was noted in both epithelial neoplastic cells (Figures 7a and 7b) and spindle stromal cells (mostly fibroblasts) (Figures 7b and 7c), supporting the presence of CSRPI expression in epithelial cells and CAFs in the scRNAseq data.



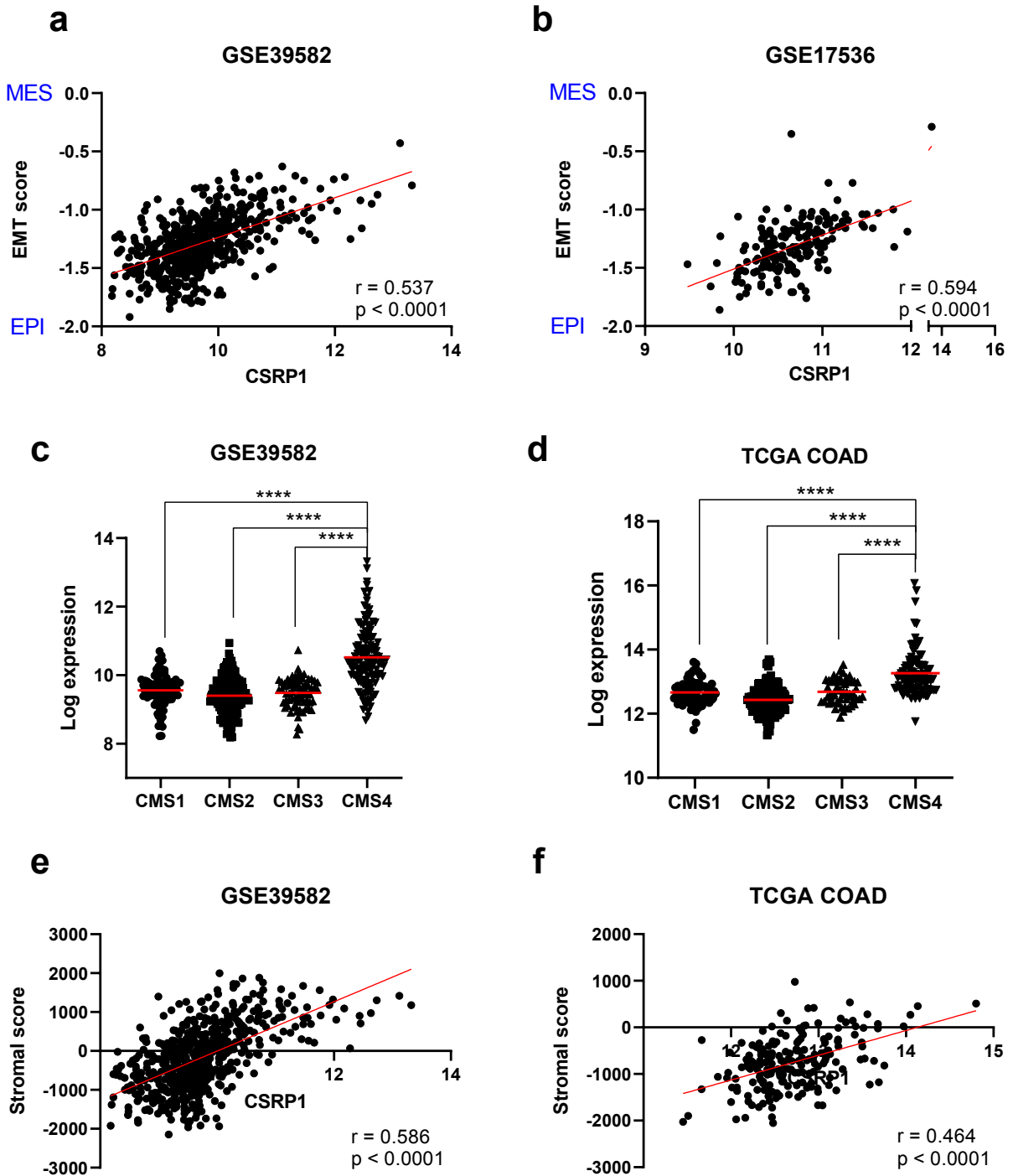
**Figure 4.** Expression of CSRP1 in colon tumors analyzed within subgroups categorized based on clinicopathological parameters. CSRP1 expression across the TNM stage (a), grade (b), and MSI status (c) and KRAS-BRAF mutation status (d) are plotted. Mutant: with KRAS or BRAF mutation, wild type: wild type for both genes. dMMR: deficient mismatch repair, pMMR: proficient mismatch repair, Mod: moderately, diff: differentiated. The GSE39582 (a,c,d) and GSE17536 (b) datasets were used. Tukey's test for multiple comparison p-value; \*  $p < 0.05$ , \*\*  $p < 0.01$ .

**Table 2.** Linear correlation of CSRP1 expression with the expression of EMT markers.

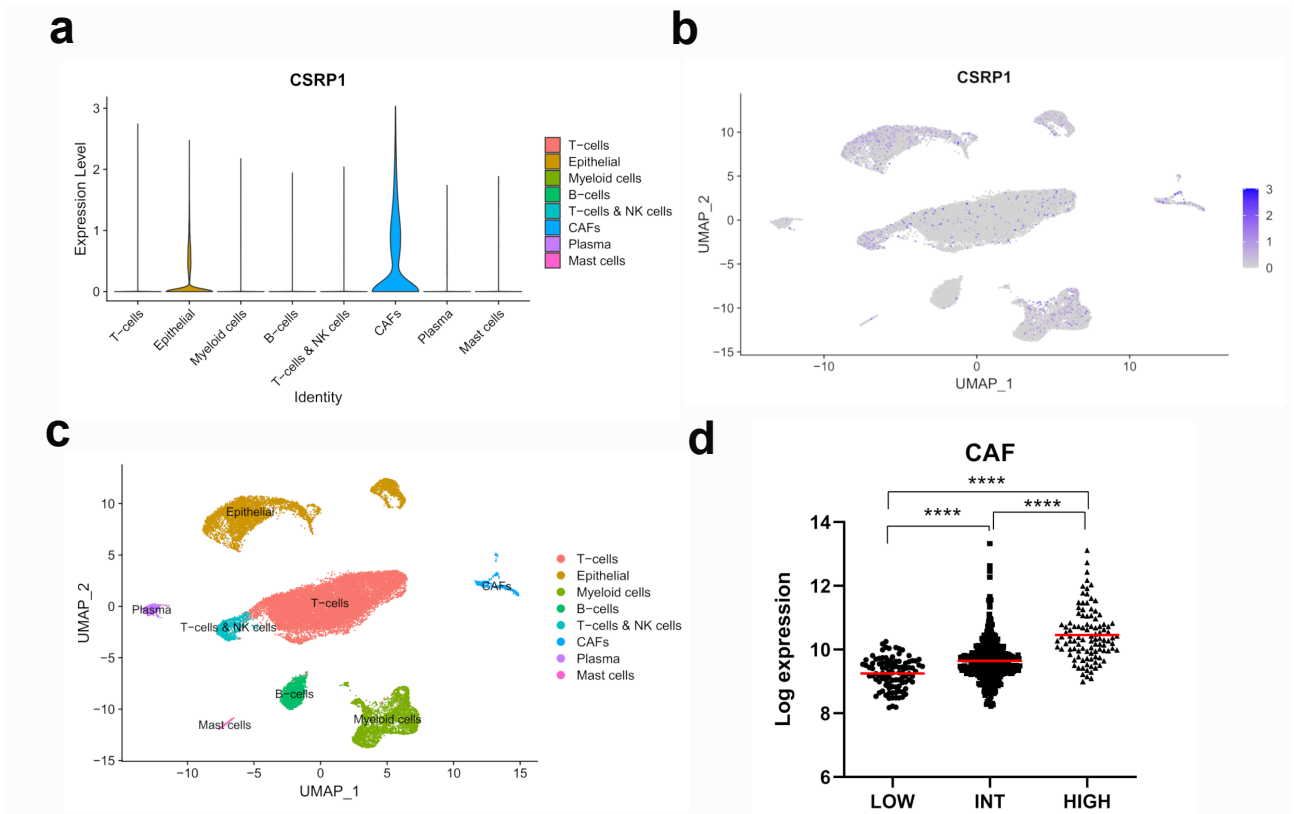
	GSE39582	TCGA COAD	Mean $r^*$
	(n = 566)	(n = 451)	
SNAI1	0.183	0.223	0.203
SNAI2	0.385	0.357	0.371
ZEB1	0.496	0.478	0.487
ZEB2	0.470	0.361	0.416
TWIST1	0.422	0.363	0.392
TWIST2	0.598	0.431	0.514
VIM	0.557	0.495	0.526

Pearson r-values are indicated. All correlations were significant ( $p < 0.001$ ). \*Average of the 2 r-values from GSE39582 and TCGA COAD.

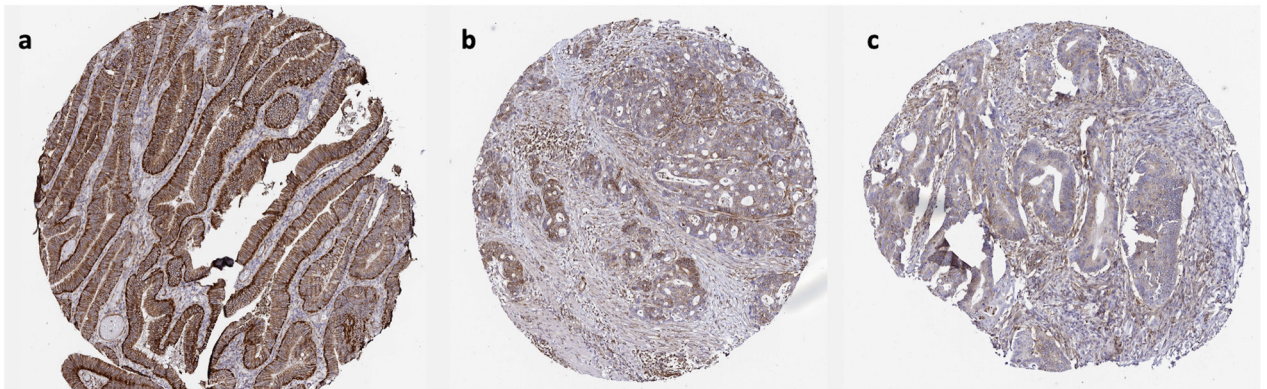




**Figure 5.** Higher CSR1 expression in CMS4, mesenchymal, and stroma-rich tumors. Correlation of CSR1 expression and EMT score is shown for the GSE39582 (a) and GSE17536 (b) datasets. EPI: epithelial, MES: mesenchymal. CSR1 expression in the CMSs for the GSE39582 (c) and TCGA COAD (d) datasets are plotted. Correlation of CSR1 and stromal score is plotted for the GSE39582 (e) and TCGA COAD (f) cohorts. Pearson r- and p-values are indicated. \*\*\*\*  $p < 0.0001$ .



**Figure 6.** CSR1 is expressed in the epithelial cells and CAFs in primary colorectal tumors. scRNAseq (GSE178318) data were processed and expression of CSR1 is shown via a violin plot for 8 cell clusters (a). Feature plots depicting the CSR1 expression (b) and cell clusters (c) are indicated. The intensity of the purple indicates the level of expression. The CSR1 expression is shown in the tumor subgroups with low, intermediate (INT), and high CAF marker expression in GSE39582 (d). \*\*\*\* Unpaired t-test  $p < 0.0001$ .



**Figure 7.** IHC-based staining for CSR1. Representative cases with CSR1 expression primarily in the epithelial neoplastic (a) and stromal (c) cells, and cases with expression in both cell types (b).

**4. Discussion**

CSR1, a smooth muscle marker that has a role in actin filament bundling, has been previously implicated in multiple cancer types. A lower expression of CSR1 in tumors compared to normal colon tissue was reported

in sporadic CRC cases by Zhou et al., and thus CSR1 was suggested as a putative tumor suppressor [10]. Consistently, the data herein showed that its expression was significantly lower in colorectal tumors compared to normal colon and rectum tissues. In HCC, the CSR1

expression was suppressed via promoter methylation [7], and, in line with that, the data herein showed a significant negative correlation between RNAseq-based expression and the methylation of the CSRPI gene in colon cancer. These data suggested that the CSRPI expression may also be regulated via promoter methylation in CRC. However, the tumor-suppressor-like expression and methylation changes of CSRPI need further investigation since its relationship with poor prognosis was shown. Although the regulation of CSRPI expression has not yet been fully elucidated, TGF- $\beta$ 1 was shown to play a role in inducing CSRPI expression through the Smad and nonconventional p38 MAPK signaling pathways [6], which may also be involved in regulating the CSRPI expression during malignant transformation.

Analysis of the prognostic relationships of CSRPI in the TCGA COAD dataset previously showed that its expression was associated with unfavorable prognosis [11]. However, this result has not yet been evaluated in independent cohorts. As reevaluation of the data in a completely different set of samples is crucial to confirm the validity of the putative molecular markers [29], the current study revealed that a high CSRPI expression was consistently associated with worse prognosis in independent cohorts. A prognostic analysis with multiple different measures of clinical outcome (OS, RFS, and DSS) further indicated that the CSRPI expression is linked with both the disease aggressiveness and time to death. The fact that these relationships were shown based on transcriptomic data generated from multiple technologies (RNAseq and microarray) further supports the robustness and power of CSRPI as a prognostic marker. The prognostic relationships of CSRPI have been studied in several different cancer types. CSRPI was included in a gene signature that identified a high-risk subgroup of acute myeloid leukemia with worse prognosis [30]. In contrast, in prostate cancer, it was reported that a high expression of CSRPI was associated with longer DFS [8]. Therefore, CSRPI may be related to different processes in cancers of different tissues of origin. Alternatively, the expression of CSRPI in the tumor microenvironment of these tissues may be a critical factor underlying its relationships with disease progression. A recent study showed that the knockdown of CSRPI in colon cancer cell lines caused a reduction in cell proliferation and migration [11]. Several studies have revealed the roles of CSRPI also in cellular contractility and movement. A reduction in cellular contractility was noted in human fibroblasts following CSRPI silencing [6]. The involvement of CSRPI in cellular movement has been shown in zebrafish, as the knockdown of CSRPI caused abnormal convergent extension cell movement,

resulting in severe deformities in midline structures [31]. Overall, these data highlight the relationship of the CSRPI expression with aggressive cell behavior in specific tissues and malignancies.

EMT is one of the major mechanisms involved in cancer progression and poor clinical outcomes [32]. Thus, linear relationships between the expression of CSRPI and EMT markers were investigated to explore the potential underlying molecular mechanisms that can explain the relation of high CSRPI expression and poor prognosis. Positive correlations between the expression of CSRPI and EMT markers were consistent when tested in independent cohorts and with different technologies measuring transcriptomic levels (microarray, RNAseq). Therefore, CSRPI may be related to cancer progression via various mechanisms related to proliferation, EMT, and cell migration.

In silico analyses of bulk transcriptomic data from colon tumors showed that tumors with a high CSRPI expression had a stroma-rich molecular profile that overlapped, to a large extent, with a CMS4 phenotype. In line with that, the data herein revealed the expression of CSRPI mainly in epithelial cells and CAFs in the tumor microenvironment. CSRPI was previously shown to be regulated by TGF- $\beta$ 1, inducing myofibroblast differentiation. Furthermore, the CSRPI expression was significantly higher in idiopathic pulmonary fibrosis compared to control lung tissue [6]. These data suggest that CSRPI may have roles in cancer progression, not only via its roles in cancer cells, but also through its potential roles in stromal cells, likely fibroblasts, in the tumor microenvironment.

In short, the CSRPI expression was associated with a more mesenchymal, stroma-rich molecular profile and poor prognosis in colon cancer. With further large-scale validation, the CSRPI gene might be shown to have the potential to contribute to patient stratification in colon cancer. Future studies investigating the role of CSRPI in CRC stroma may pave the way to the evaluation of this gene as a potential therapeutic target.

#### **Acknowledgment**

The author thanks Meral Üner of the Department of Pathology, Hacettepe University, Türkiye, for her valuable guidance regarding the interpretation of the Human Protein Atlas data.

#### **Informed consent**

Not applicable. Datasets used in this study were obtained from public databases, as described in the methods section.

## References

- Pomies P, Louis HA, Beckerle MC. CRP1 a LIM domain protein implicated in muscle differentiation, interacts with alpha-actinin. *The Journal of Cell Biology* 1997; 139 (1): 157-168. <https://doi.org/10.1083/jcb.139.1.157>
- Henderson JR, Macalma T, Brown D, Richardson JA, Olson EN et al. The LIM protein, CRP1, is a smooth muscle marker. *Developmental Dynamics* 1999; 214 (3): 229-238. [https://doi.org/10.1002/\(SICI\)1097-0177\(199903\)214:3<229::AID-AJA6>3.0.CO;2-S](https://doi.org/10.1002/(SICI)1097-0177(199903)214:3<229::AID-AJA6>3.0.CO;2-S)
- Tran TC, Singleton C, Fraley TS, Greenwood JA. Cysteine-rich protein 1 (CRP1) regulates actin filament bundling. *BioMed Central Cell Biology* 2005; 6: 45. <https://doi.org/10.1186/1471-2121-6-45>
- Dubé JY, Chapdelaine P, Trahan PL, Deperthes D, Frenette G et al. Abundant cysteine-rich protein-1 is localized in the stromal compartment of the human prostate. *Archives of Andrology* 1998; 40 (2): 109-115. <https://doi.org/10.3109/01485019808987933>
- Latonen L, Järvinen PM, Laiho M. Cytoskeleton-interacting LIM-domain protein CRP1 suppresses cell proliferation and protects from stress-induced cell death. *Experimental Cell Research* 2008; 314 (4): 738-747. <https://doi.org/10.1016/j.yexcr.2007.11.024>
- Järvinen PM, Myllärniemi M, Liu H, Moore HM, Leppäranta O et al. Cysteine-rich protein 1 is regulated by transforming growth factor-beta1 and expressed in lung fibrosis. *Journal of Cellular Physiology* 2012; 227 (6): 2605-2612. <https://doi.org/10.1002/jcp.23000>
- Hirasawa Y, Arai M, Imazeki F, Tada M, Mikata R et al. Methylation status of genes upregulated by demethylating agent 5-aza-2'-deoxycytidine in hepatocellular carcinoma. *Oncology* 2006; 71 (1-2): 77-85. <https://doi.org/10.1159/000100475>
- Chen X, Ma J, Xu C, Wang L, Yao Y et al. Identification of hub genes predicting the development of prostate cancer from benign prostate hyperplasia and analyzing their clinical value in prostate cancer by bioinformatic analysis. *Discover Oncology* 2022; 13 (1): 54. <https://doi.org/10.1007/s12672-022-00508-y>
- Gao L, Ye M, Lu X, Huang D. Hybrid method based on information gain and support vector machine for gene selection in cancer classification. *Genomics Proteomics Bioinformatics* 2017; 15 (6): 389-395. <https://doi.org/10.1016/j.gpb.2017.08.002>
- Zhou CZ, Qiu GQ, Wang XL, Fan JW, Tang HM et al. Screening of tumor suppressor genes on 1q31.1-32.1 in Chinese patients with sporadic colorectal cancer. *Chinese Medical Journal* 2008; 121 (24): 2479-2486
- Yu S, Zhao H, Meng H, Shi S, Cao S et al. CSRP1 Promotes colon adenocarcinoma growth and serves as an independent risk biomarker for worse prognosis. *Genetics Research* 2023; 2023: 8586507. <https://doi.org/10.1155/2023/8586507>
- Marisa L, de Reyniès A, Duval A, Selves J, Gaub MP et al. Gene expression classification of colon cancer into molecular subtypes: characterization, validation, and prognostic value. *PLoS Medicine* 2013; 10 (5): e1001453. <https://doi.org/10.1371/journal.pmed.1001453>
- Smith JJ, Deane NG, Wu F, Merchant NB, Zhang B et al. Experimentally derived metastasis gene expression profile predicts recurrence and death in patients with colon cancer. *Gastroenterology* 2010; 138 (3): 958-968. <https://doi.org/10.1053/j.gastro.2009.11.005>
- Demirkol Canlı S, Seza EG, Sheraj I, Gönçeli I, Turhan N et al. Evaluation of an aldo-keto reductase gene signature with prognostic significance in colon cancer via activation of epithelial to mesenchymal transition and the p70S6K pathway. *Carcinogenesis* 2020; 41 (9): 1219-1228. <https://doi.org/10.1093/carcin/bgaa072>
- Yoshihara K, Shahmoradgoli M, Martínez E, Vegesna R, Kim H et al. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nature Communications* 2013; 4: 2612. <https://doi.org/10.1038/ncomms3612>
- Demirkol Canlı S. Evaluation of prognostic markers in cancer-associated fibroblast based sub-groups of colorectal cancer. *Acta Medica* 2022; 53 (2): 133-143. <https://doi.org/10.32552/2022.ActaMedica.682>
- Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* 2014; 15 (12): 550-571. <https://doi.org/10.1186/s13059-014-0550-8>
- Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discovery* 2012; 2 (5): 401-404. <https://doi.org/10.1158/2159-8290.CD-12-0095>
- Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Science Signaling* 2013; 6 (269): pii. <https://doi.org/10.1126/scisignal.2004088>
- Tang Z, Li C, Kang B, Gao G, Li C et al. GEPIA: A web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Research* 2017; 45 (W1): W98-W102. <https://doi.org/10.1093/nar/gkx247>
- Demirkol S, Gomçeli I, Isbilen M, Dayanc BE, Tez M et al. A combined ULBP2 and SEMA5A expression signature as a prognostic and predictive biomarker for colon cancer. *Journal of Cancer* 2017; 8 (7): 1113-1122. <https://doi.org/10.7150/jca.17872>
- Satija R, Farrell JA, Gennert D, Schier AF, Regev A. Spatial reconstruction of single-cell gene expression data. *Nature Biotechnology* 2015; 33 (5): 495-502. <https://doi.org/10.1038/nbt.3192>
- McGinnis CS, Murrow LM, Gartner ZJ. DoubletFinder: doublet detection in single-cell RNA sequencing data using artificial nearest neighbors. *Cell Systems* 2019; 8 (4): 329-337 e4. <https://doi.org/10.1016/j.cels.2019.03.003>

24. Che LH, Liu JW, Huo JP, Luo R, Xu RM et al. A single-cell atlas of liver metastases of colorectal cancer reveals reprogramming of the tumor microenvironment in response to preoperative chemotherapy. *Cell Discovery* 2021; 7 (1): 80. <https://doi.org/10.1038/s41421-021-00312-y>
25. Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P et al. Proteomics. Tissue-based map of the human proteome. *Science* 2015; 347 (6220): 1260419. <https://doi.org/10.1126/science.1260419>
26. Pontén F, Jirstrom K, Uhlen M. The Human Protein Atlas--a tool for pathology. *The Journal of Pathology* 2008; 216 (4): 387-393. <https://doi.org/10.1002/path.2440>
27. Sabouni E, Nejad MM, Mojtavavi S, Khoshduz S, Mojtavavi M et al. Unraveling the function of epithelial-mesenchymal transition (EMT) in colorectal cancer: Metastasis, therapy response, and revisiting molecular pathways. *Biomedicine Pharmacotherapy* 2023; 160: 114395. <https://doi.org/10.1016/j.biopha.2023.114395>
28. Guinney J, Dienstmann R, Wang X, de Reyniès A, Schlicker A et al. The consensus molecular subtypes of colorectal cancer. *Nature Medicine* 2015; 21 (11): 1350-1356. <https://doi.org/10.1038/nm.3967>
29. Henry NL and Hayes DF. Cancer biomarkers. *Molecular Oncology* 2012; 6 (2): 140-146. <https://doi.org/10.1016/j.molonc.2012.01.010>
30. Zhao C, Wang Y, Sharma A, Wang Z, Zheng C et al. Identification of the integrated prognostic signature associated with immuno-relevant genes and long non-coding RNAs in acute myeloid leukemia. *Cancer Investigation* 2022; 40 (8): 663-674. <https://doi.org/10.1080/07357907.2022.2096230>
31. Miyasaka KY, Kida YS, Sato T, Minami M, Ogura T. Csrp1 regulates dynamic cell movements of the mesendoderm and cardiac mesoderm through interactions with dishevelled and diversin. *Proceedings of the National Academy of Sciences of the United States of America* 2007; 104 (27): 11274-11279. <https://doi.org/10.1073/pnas.0702000104>
32. Yang J, Weinberg RA. Weinberg, Epithelial-mesenchymal transition: at the crossroads of development and tumor metastasis. *Development Cell* 2008; 14 (6): 818-829. <https://doi.org/10.1016/j.devcel.2008.05.009>