

Research Article

An Integrated Bioinformatic Analysis of the S100 Gene Family for the Prognosis of Colorectal Cancer

Meng-Lu Zeng ¹, Xian-Jin Zhu ¹, Jin Liu ¹, Peng-Chong Shi ¹, Yan-Li Kang,²
Zhen Lin ¹ and Ying-Ping Cao ¹

¹Department of Clinical Laboratory Medicine, Fujian Medical University Union Hospital, Fuzhou 350000, China

²Department of Clinical Laboratory Medicine, Fujian Provincial Hospital, Fuzhou 350001, China

Correspondence should be addressed to Ying-Ping Cao; caoyingping918@hotmail.com

Meng-Lu Zeng and Xian-Jin Zhu contributed equally to this work.

Received 11 July 2020; Accepted 28 October 2020; Published 27 November 2020

Academic Editor: Alessandro De Vita

Copyright © 2020 Meng-Lu Zeng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Background. S100 family genes exclusively encode at least 20 calcium-binding proteins, which possess a wide spectrum of intracellular and extracellular functions in vertebrates. Multiple lines of evidences suggest that dysregulated S100 proteins are associated with human malignancies including colorectal cancer (CRC). However, the diverse expression patterns and prognostic roles of distinct S100 genes in CRC have not been fully elucidated. **Methods.** In the current study, we analyzed the mRNA expression levels of S100 family genes and proteins and their associations with the survival of CRC patients using the OncoPrint analysis and GEPIA databases. Expressions and mutations of S100 family genes were analyzed using the cBioPortal, and protein-protein interaction (PPI) networks of S100 proteins and their mutation-related coexpressed genes were analyzed using STRING and Cytoscape. **Results.** We observed that the mRNA expression levels of S100A2, S100A3, S100A9, S100A11, and S100P were higher and the level of S100B was lower in CRC tissues than those in normal colon mucosa. A high S100A10 levels was associated with advanced-stage CRC. Results from GEPIA database showed that highly expressed S100A1 was correlated with worse overall survival (OS) and disease-free survival (DFS) and that overexpressions of S100A2 and S100A11 were associated with poor DFS of CRC, indicating that S100A1, S100A2, and S100A11 are potential prognostic markers. Unexpectedly, most of S100 family genes showed no significant prognostic values in CRC. **Conclusions.** Our findings, though still need to be ascertained, offer novel insights into the prognostic implications of the S100 family in CRC and will inspire more clinical trials to explore potential S100-targeted inhibitors for the treatment of CRC.

1. Introduction

The S100 family, with a common Ca²⁺-binding motif, EF-hand, contains a group of low molecular weight acidic polypeptides (M_r between 9 and 14kDa) [1], of which more than 20 S100 proteins are encoded in the human genome [2]. The S100 proteins appear to be involved in a multitude of biological processes, including calcium homeostasis, cell growth, invasion and motility, apoptosis, protein phosphorylation, chemotaxis, and inflammation [1–3]. Extensive evidence suggests that the deregulated expression of S100 proteins is closely linked to tumor progression and drug

resistance in the treatment of many malignant tumors, including ovarian cancer [4], breast cancer [5], prostate cancer [6], and colorectal cancer [7].

Colorectal cancer (CRC), with high morbidity and mortality, is one of the most common malignant cancers of the digestive tract worldwide [8]. The prevalence of CRC has gradually increased owing to environmental deterioration and unhealthy lifestyle, as well as the contribution of new diagnostic techniques [9]. Despite considerable improvements in the diagnosis and treatment of CRC, many patients are diagnosed at advanced stages or relapse, which is associated with a poor prognosis for survival. A previous study

showed that the 5-year overall survival (OS) rate of patients with metastatic colorectal cancer (mCRC) remains less than 15% [10]. Hence, identifying biomarkers for diagnosis and prognosis of CRC is the first imperative for developing valuable prognostic markers and individualized therapeutics.

Some relevant literatures report abnormal expressions of the S100 genes and their associations with clinicopathological characteristics and prognosis in human CRC. To the best of the authors' knowledge, the roles of S100s in CRC have not yet been explored using bioinformatics analysis. Integrating the online high-throughput microarray analysis of gene expression and copy number variants (CNVs) from massive platform data, we exhaustively analyzed the expressions and mutations of various S100 genes to determine the distinct expression patterns, numerous functions, and potential prognostic value of S100s in CRC.

2. Materials and Methods

2.1. Oncomine Analysis. Oncomine (<http://www.oncomine.org>), an online gene expression array database and web-based data-mining platform containing 715 datasets and 86733 samples, is frequently used to stimulate discovery in genome-wide expression analyses. Here, the mRNA levels of S100s were analyzed by Oncomine in different cancers. The expression levels of S100s were compared between clinical cancer specimens and normal controls by performing Student's *t*-test and assessing the *p* value. The fold change was set as 2 and the threshold of the *p* value was set as 0.01. The other parameters of Oncomine were set as the default settings.

2.2. Gene Expression Profiling Interactive Analysis (GEPIA) Dataset. GEPIA, available at <http://gepia.cancer-pku.cn/> (March 11, 2020), is a database that provides diverse functions, including tumor and normal differential expression analysis, correlation analysis, profiling plotting, patient survival analysis, dimensionality reduction analysis, and the detection of similar genes based on different human tumor and normal samples from The Cancer Genome Atlas (TCGA) and Genotype-Tissue Expression (GTEx) programs [11]. The profiling, boxplot, and stage plot expressions of the S100 family genes were obtained from colon adenocarcinoma (COAD) and rectal adenocarcinoma (READ) tissue order via GEPIA, and the other default parameters of GEPIA were reserved.

The prognostic values of the mRNA expression of S100s were also evaluated by GEPIA. Patients with COAD and READ were split into two groups according to the median expression level (high vs. low expression) to analyze OS and disease-free survival (DFS). Patient samples were evaluated by the GEPIA survival plot, and the hazard ratio (HR) was presented with a 95% confidence interval (CI) and log *p* value.

2.3. cBioportal Analysis with TCGA Data. The TCGA database is a publicly funded project that includes high-throughput sequencing and pathological data of over 30 different human cancers [12]. cBioPortal (<http://www.cbioportal.org/>; accessed

March 11, 2020) was applied to analyze S100s from the COAD (TCGA, Firehose Legacy) dataset including 379 cases. Mutations and putative CNAs from Genomic Identification of Significant Targets in Cancer (GISTIC) were selected as genomic profiles. Moreover, mRNA expression Z-scores relative to diploid samples (RNA Seq V2 RSEM) were chosen for the mRNA expression of genomic profiles, and protein expression Z-scores (RPPA) were selected for the protein/phosphoprotein expression levels. In addition, the top 20 coexpressed genes of the S100 family were also calculated for further analysis based on cBioPortal's online instructions according to the *p* value.

2.4. Integration of the Protein-Protein Interaction (PPI) Network and Module Analysis. The Search Tool for the Retrieval of Interacting Genes/Proteins (STRING, <https://string-db.org/>, version 11.0; accessed March 11, 2020) is an online database designed to predict PPI network information [13]. In the present study, the coexpression PPI network of the S100 family genes was analyzed via the STRING database, and 20 coexpressed genes with a combined score > 0.4 were relatively significantly altered. The PPI network of those genes mentioned above was constructed and visualized by the Cytoscape software (version 3.6.1), and the Molecular Complex Detection (MCODE) plug-in application of the software was used to find important modules for analysis with the criterion set as follows: *k*-core = 2, node score cutoff = 0.2, degree cutoff = 2, and maximum depth = 100.

2.5. Gene Ontology (GO) and Pathway Enrichment Analysis. GO analysis, covering the molecular function (MF), cellular component (CC), and biological process (BP) categories, is a commonly used method to study the characteristic attributes of large-scale genomic and transcriptome data [14]. Kyoto Encyclopedia of Genes and Genomes (KEGG) is a systematic collection of online web servers providing gene function and biological pathway information [15]. The Database for Annotation, Visualization, and Integrated Discovery (DAVID, <https://david.ncicrf.gov/>; version 6.8; accessed March 11, 2020) is a free online bioinformatic resource that is designed to provide exhaustive functional annotation tools to identify enriched GO terms and visualize genes on KEGG pathway maps. The S100 family genes and their coexpressed genes of the top 5 modules in Cytoscape were input into the DAVID online tools to obtain the GO functions and KEGG pathways. Terms with a *p* value < 0.05 were considered statistically significant.

3. Results

3.1. Expression Levels of the S100 Gene Family in Patients with CRC. We imported 21 genes in the S100 family reported by Anne et al. in the Oncomine database and compare their expression levels between the normal and cancerous CRC samples (see Figure 1). Using datasets from Oncomine, the mRNA expression levels of S100A2, S100A3, S100A6, S100A8, S100A9, S100A11, and S100P were significantly upregulated (a fold change of >1.5) in CRC tissues. Among various pathological types of CRC, two common types

TABLE 1: The changes in the expression of the S100 family genes at the transcription level between different types of colorectal cancer and colon tissues (Oncomine database).

	Types of colorectal cancer vs. normal colon	Fold change	<i>p</i> value	<i>t</i> -test	Ref
S100A1	Rectal adenocarcinoma vs. normal	1.009	0.457	0.111	Kaiser Colon [25]
	Colon adenocarcinoma vs. normal	1.003	0.484	0.043	Kaiser Colon [25]
	Colorectal adenocarcinoma vs. normal	1.042	0.143	1.077	Skrzypczak Colorectal [27]
S100A2	Colorectal carcinoma vs. Normal	5.846	2.66E-13	9.915	Skrzypczak Colorectal [27]
	Colon adenocarcinoma vs. normal	2.591	4.94E-07	7.205	Kaiser Colon [25]
	Rectal adenocarcinoma vs. normal	9.983	5.77E-46	24.76	Gaedcke Colorectal [26]
S100A3	Colorectal carcinoma vs. normal	1.6	1.82E-08	6.425	Skrzypczak Colorectal [27]
	Rectal adenocarcinoma vs. normal	2.803	4.20E-24	14.11	Gaedcke Colorectal [26]
	Colorectal carcinoma vs. normal	1.73	1.18E-04	3.923	Skrzypczak Colorectal [27]
S100A4	Colon adenocarcinoma vs. normal	1.292	8.00E-02	1.52	Kaiser Colon [25]
	Rectal adenocarcinoma vs. normal	1.391	1.00E-01	1.365	Kaiser Colon [25]
	Colorectal carcinoma vs. normal	1.066	7.00E-02	1.498	Skrzypczak Colorectal [27]
S100A5	Colon adenocarcinoma vs. normal	1.09	2.70E-02	2.212	Kaiser Colon [25]
	Rectal adenocarcinoma vs. normal	1.057	1.57E-01	1.055	Kaiser Colon [25]
S100A6	Colon adenocarcinoma vs. normal	2.037	6.88E-04	4.931	Kaiser Colon [25]
	Rectal adenocarcinoma vs. normal	2.735	5.54E-04	4.409	Kaiser Colon [25]
	Colorectal carcinoma vs. normal	1.411	6.00E-03	2.637	Skrzypczak Colorectal [27]
S100A7	Colon adenocarcinoma vs. normal	1.293	3.52E-04	3.777	Kaiser Colon [25]
	Rectal adenocarcinoma vs. normal	1.27	5.00E-03	3.11	Kaiser Colon [25]
	Colon adenocarcinoma vs. normal	1.875	2.70E-04	7.056	Kaiser Colon [25]
S100A7A	Rectal adenocarcinoma vs. normal	1.987	1.25E-04	5.321	Kaiser Colon [25]
	NA	NA	NA	NA	NA
S100A2L2	Colorectal carcinoma vs. normal	6.313	5.37E-06	5.126	Skrzypczak Colorectal [27]
	Colon adenocarcinoma vs. normal	2.29	3.89E-04	4.478	Kaiser Colon [25]
	Rectal adenocarcinoma vs. normal	2.723	1.80E-02	2.488	Kaiser Colon [25]
S100A8	Colorectal carcinoma vs. normal	3.941	1.14E-08	6.784	Skrzypczak Colorectal [27]
	Colon adenocarcinoma vs. normal	1.951	6.05E-04	5.139	Kaiser Colon [25]
	Rectal adenocarcinoma vs. normal	2.097	1.80E-02	2.455	Kaiser Colon [25]
S100A9	Colorectal adenocarcinoma vs. normal	-1.169	9.97E-01	-2.95	Skrzypczak Colorectal [27]
	Colon adenocarcinoma vs. normal	-1.075	7.60E-01	-0.719	Kaiser Colon [25]
	Rectal adenocarcinoma vs. normal	-1.107	6.66E-01	-0.444	Kaiser Colon [25]
S100A10	Colorectal adenocarcinoma vs. normal	2.282	2.21E-11	10.392	Skrzypczak Colorectal [27]
	Colon adenocarcinoma vs. normal	1.734	4.00E-03	4.425	Kaiser Colon [25]
	Rectal adenocarcinoma vs. normal	1.775	3.00E-03	4.487	Kaiser Colon [25]
S100A11	Colorectal carcinoma vs. normal	1.695	4.00E-03	2.743	Skrzypczak Colorectal [27]
	Colon adenocarcinoma vs. normal	1.103	1.97E-01	0.856	Ki Colon [28]
	Colorectal carcinoma vs. normal	1.199	4.60E-02	1.721	Skrzypczak Colorectal [27]
S100A12	Colon adenocarcinoma vs. normal	1.133	1.51E-01	1.102	Kaiser Colon [25]
	Rectal adenocarcinoma vs. normal	1.081	3.13E-01	0.501	Kaiser Colon [25]
	Colorectal carcinoma vs. normal	-1.655	1.00E+00	-3.731	Skrzypczak Colorectal [27]
S100A13	Colon adenocarcinoma vs. normal	-2.18	1.00E+00	-8.84	Kaiser Colon [25]
	Rectal adenocarcinoma vs. normal	-2.143	1.00E+00	-5.957	Kaiser Colon [25]
	Colorectal carcinoma vs. normal	-1.573	9.92E-01	-3.348	Kaiser Colon [25]
S100A14	Colon adenocarcinoma vs. normal	-1.549	9.91E-01	-3.084	Kaiser Colon [25]
	Colorectal carcinoma vs. normal	-1.065	6.46E-01	-0.378	Skrzypczak Colorectal [27]
	Colorectal carcinoma vs. normal	3.212	1.91E-06	5.587	Skrzypczak Colorectal [27]
S100B	Colon adenocarcinoma vs. normal	4.911	3.00E-03	4.798	Kaiser Colon [25]
	Rectal adenocarcinoma vs. normal	4.879	2.00E-03	4.233	Kaiser Colon [25]

TABLE 1: Continued.

	Types of colorectal cancer vs. normal colon	Fold change	<i>p</i> value	<i>t</i> -test	Ref
S100G	Colorectal adenocarcinoma vs. normal	1.025	2.08E-01	0.821	Skrzypczak Colorectal [27]
	Colon adenocarcinoma vs. normal	1.19	1.20E-01	1.202	Notterman Colon [29]
	Rectal adenocarcinoma vs. normal	1.025	1.10E-02	2.325	Gaedcke Colorectal [26]
S100Z	Colorectal carcinoma vs. normal	1.059	4.50E-02	1.738	Skrzypczak Colorectal [27]
	Colon adenocarcinoma vs. normal	1.138	0.081	1.698	Kaiser Colon [25]
	Rectal adenocarcinoma vs. normal	1.115	1.13E-01	1.398	Kaiser Colon [25]

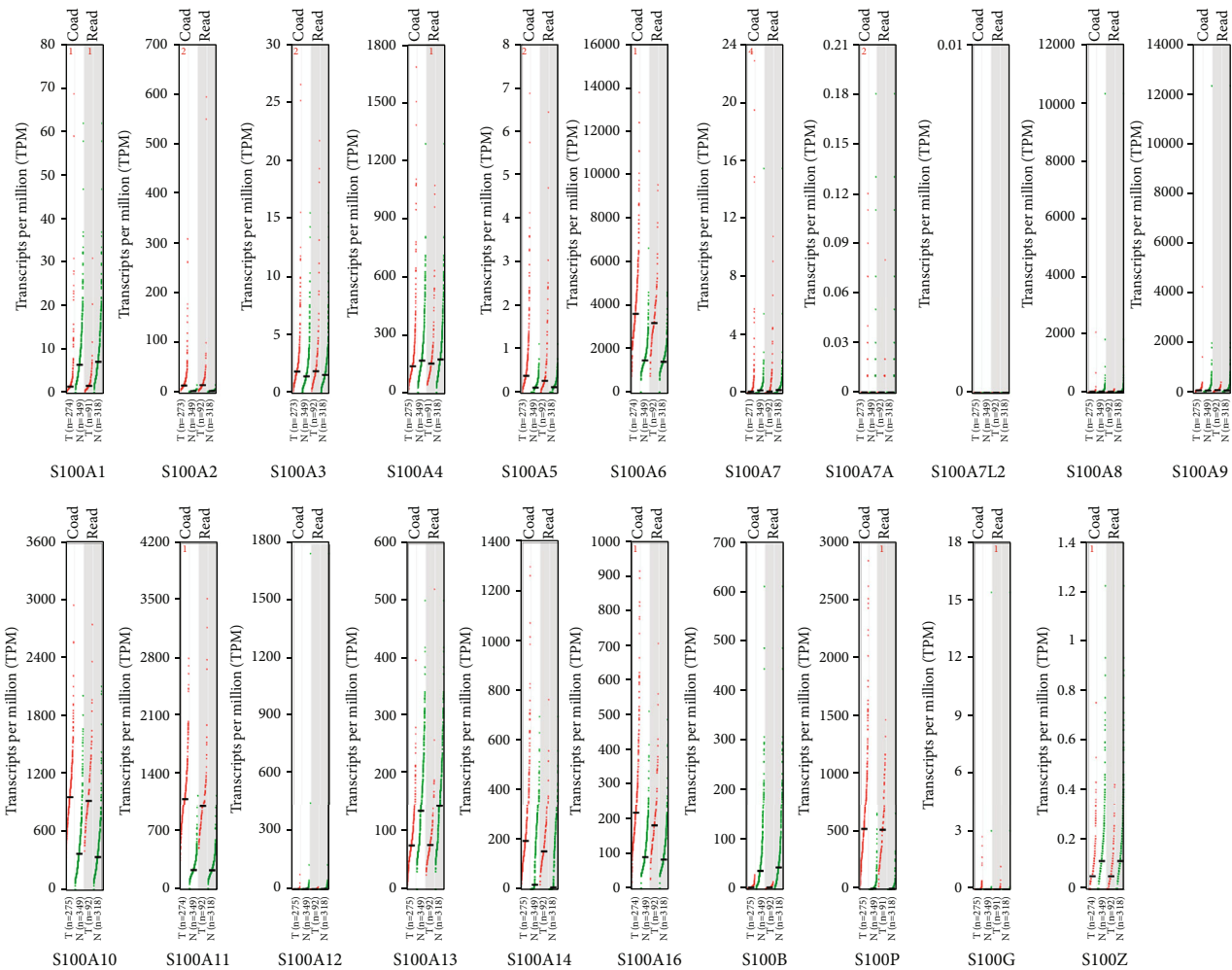


FIGURE 2: The transcriptional expression of the S100 family members in COAD and READ. Red indicates expression in tumor tissues, and green indicates expression in corresponding normal tissues.

than the levels of the other S100 genes in CRC tissues (see Figure 3). However, the detection of several genes (e.g., S100A7A, S100A7L2, and S100G) using GEPIA was unavailable due to insufficient data. The expression levels of the S100 family genes were also analyzed in COAD and READ at different stages. Only the S100A10 subgroup showed significant differences in expression levels between different stages (see Figure 4).

3.3. Survival Analysis of the S100 Gene Family in Patients with CRC by GEPIA.

Survival-associated S100 genes were

identified by GEPIA database. Despite unavailable analysis for the correlations between S100A7A, S100A7L2, and S100G expressions and OS or DFS of CRC patients due to insufficient data, the analysis for the other S100 genes revealed that the S100A1 overexpression was associated with worse OS of patients with COAD and READ ($p < 0.05$) (see Figure 5) and that S100A2–S100A14, S100A16, S100B, S100P, and S100Z expressions had no significant correlations with OS of the patients. The S100A4 overexpression was correlated with poor OS of the patients, while elevated S100A10 and S100P levels were correlated with favorable OS. It was

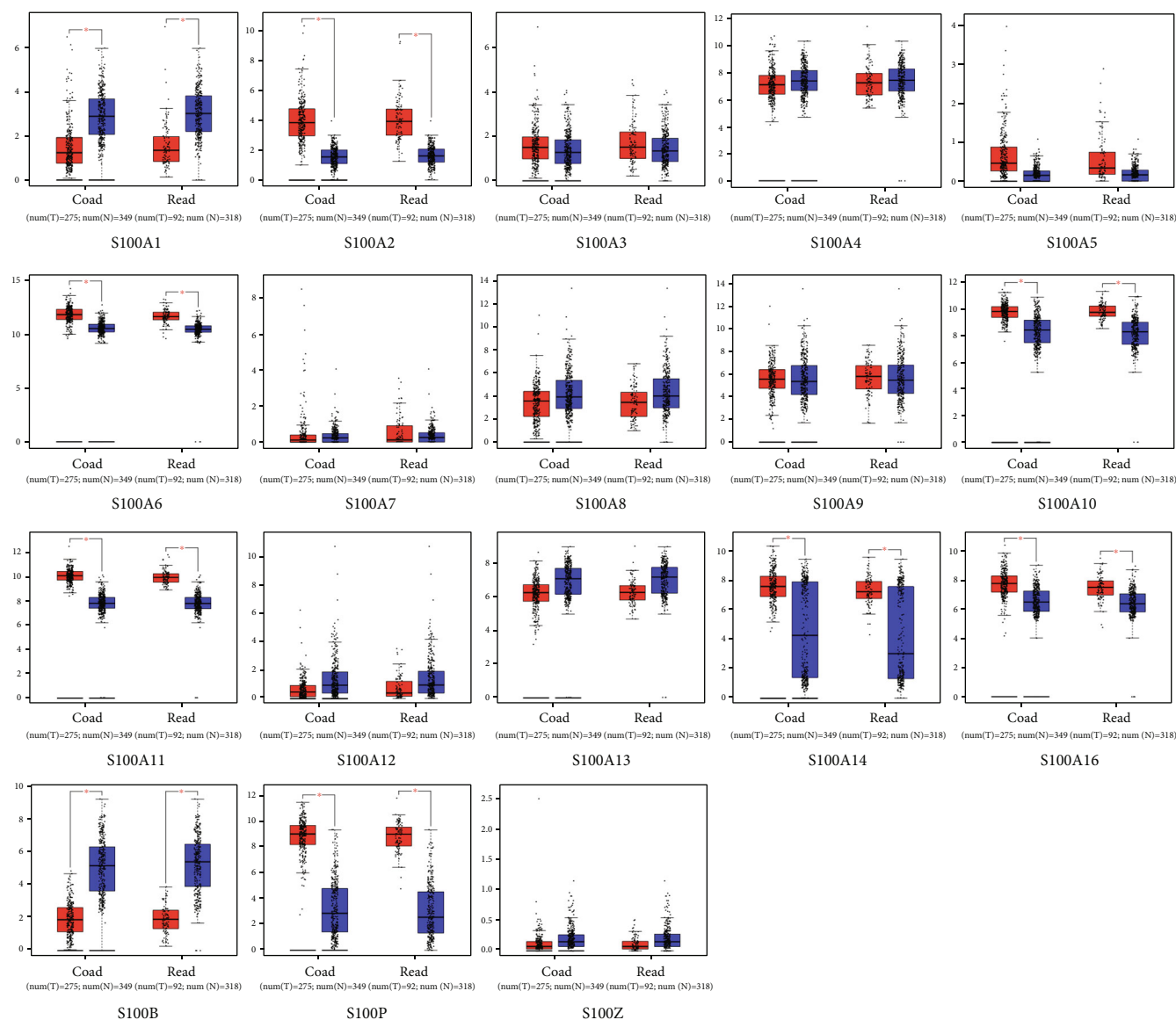


FIGURE 3: The expression of S100 family members of CRC patients. Red indicates expression in tumor tissues, and blue indicates expression in normal colon tissues. Significantly expressed genes are listed with an asterisk ($*p < 0.05$).

also found that increased S100A1, S100A2, and S100A11 mRNA levels were apparently associated with poor DFS of COAD patients (see Figure 5). An elevated S100A13 level was seemingly associated with worse DFS of COAD and READ patients ($p = 0.065$), but with a nonsignificant difference. Overexpressed S100A3 and S100Z were associated with poor DFS, while S100A7 and S100P overexpressions were associated with favorable DFS of COAD and READ patients. The other genes of the S100 family had no clear correlations with DFS in CRC.

3.4. The Correlations between S100 Family Genes in Patients with COAD. We used the cBioPortal online tool to analyze correlations among altered S100 family genes in COAD specimens and pinpointed 126 specimens (126/379, 33%) showing abnormally expressed S100 genes related to COAD, of which 31 to 69 samples exhibited two or more abnormally

expressed S100 genes (see Figure 6(a)). The Pearson correlation between these S100 genes in COAD specimens was also calculated by analyzing their mRNA expression data (RNA Seq V2 RSEM) from TCGA database (Firehose Legacy) using the cBioPortal platform. The results revealed significantly positive associations between the following pairs of S100 genes: S100A4 with S100A13; S100A6 with S100A10, S100A11, S100A13, S100A14, S100A16, and S100P; S100A7 with S100A8 and S100A9; S100A8 with S100A7, S100A9, and S100A12; S100A9 with S100A7, S100A8, and S100A12; S100A10 with S100A6, S100A11, S100A13, S100A14, S100A16, and S100P; S100A11 with S100A6, S100A10, S100A13, and S100A16; S100A12 with S100A8 and S100A9; S100A13 with S100A4, S100A6, S100A10, S100A11, S100A16, and S100P; S100A14 with S100A6, S100A10, S100A16, and S100P; S100A16 with S100A6, S100A10, S100A11, S100A13, S100A14, and S100P; and

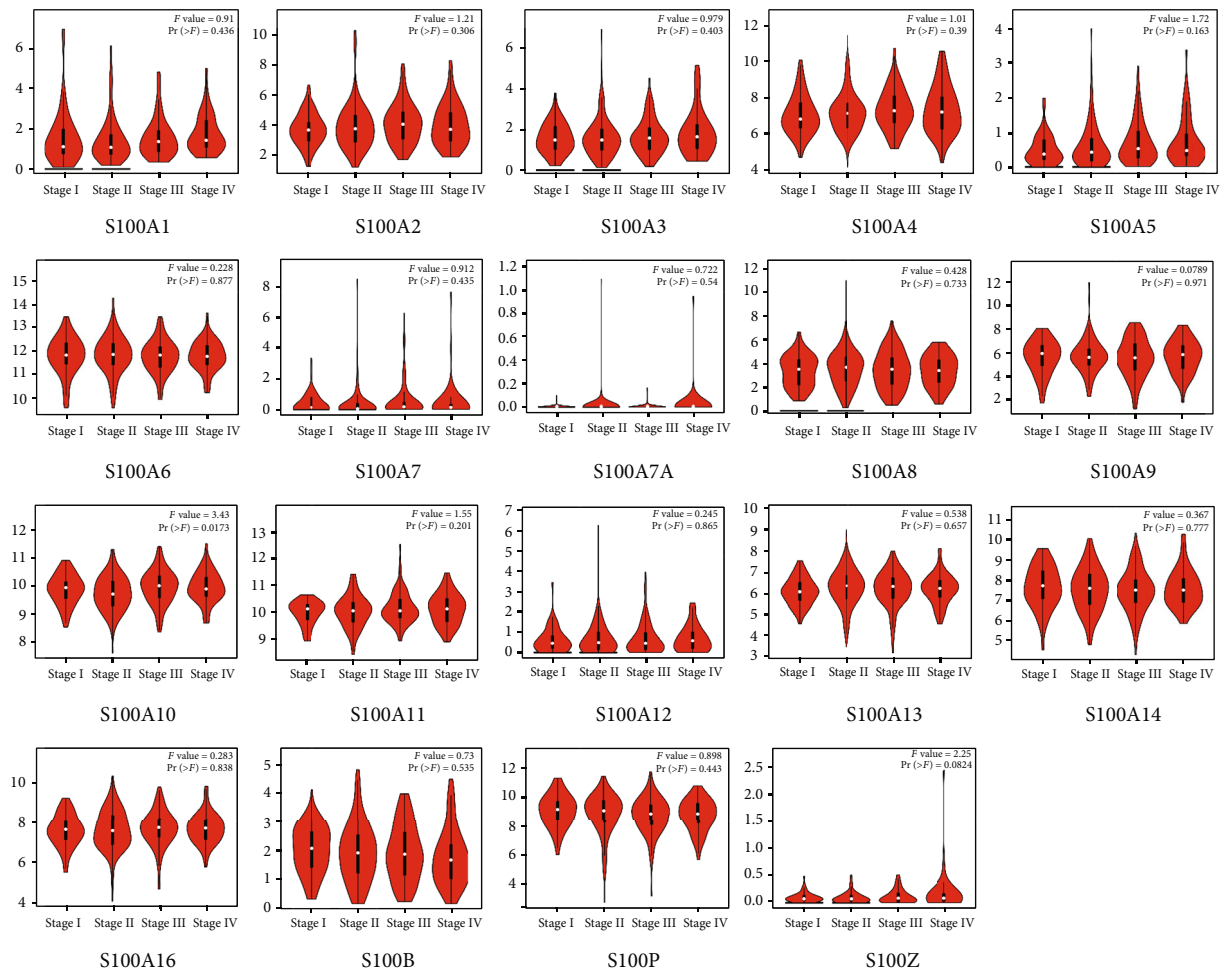


FIGURE 4: Expression of the S100 family members in CRC patients with various clinical stages.

S100P with S100A6, S100A10, S100A13, S100A14, and S100A16. Their coexpression networks were depicted in Figure 6(b).

3.5. PPI Network and Module Analysis of S100 Proteins and Their Coexpressed Genes. We then constructed a coexpressed gene network of the S100 family linked to CRC based on 20 most relevant genes. The PPI network of those genes mentioned above was obtained from STRING, and the results were visualized by the Cytoscape software. A network of 212 nodes and 574 edges was constructed, and 281 genes were analyzed through the MCODE plug-in (see Figure 7(a)). Additionally, the top 5 modules with default parameters were selected to elucidate the interactions between S100 proteins and other molecules. Finally, MCODE analysis generated 5 modules as follows: Module 1 with 12 nodes and 66 edges (see Figure 7(b)), Module 2 with 9 nodes and 21 edges (see Figure 7(c)), Module 4 with 8 nodes and 13 edges (see Figure 7(d)), Module 3 with 12 nodes and 27 edges (see Figure 7(e)), and Module 5 with 11 nodes and 18 edges. S100 family genes were mostly distributed among the top 5 modules.

3.6. GO Function and KEGG Pathway Enrichment Analyses of CRC-Related S100 Genes. The enriched GO functions of S100 family genes and the 5 modules of coexpressed genes were analyzed by DAVID online database for the CC, MF, and BP categories. As shown in Table 2, the top 6 GO terms of CC for these coexpressed genes consisted of extracellular exosomes, extracellular regions, the cornified envelope, the plasma membrane, the perinuclear region of cytoplasm, and an integral component of the plasma membrane. The top 6 GO terms of MF involved RAGE receptor binding, calcium ion binding, S100 protein binding, calcium-dependent protein binding, interleukin-8 binding, and Toll-like receptor 4 binding. The top 6 GO terms of BP included inflammatory responses, chemotaxis, peptide cross-linking, chemokine-mediated signaling pathways, G-protein coupled receptor signaling pathways, and positive regulation of cytosolic calcium ion concentration. The KEGG pathway enrichment analyses (see Table 3) showed that S100 alterations and the coexpressed genes that altered frequently were particularly enriched in the chemokine signaling pathways ($p = 0.001459$) and cytokine-cytokine receptor interaction ($p = 0.003873$) (see Figure 8).

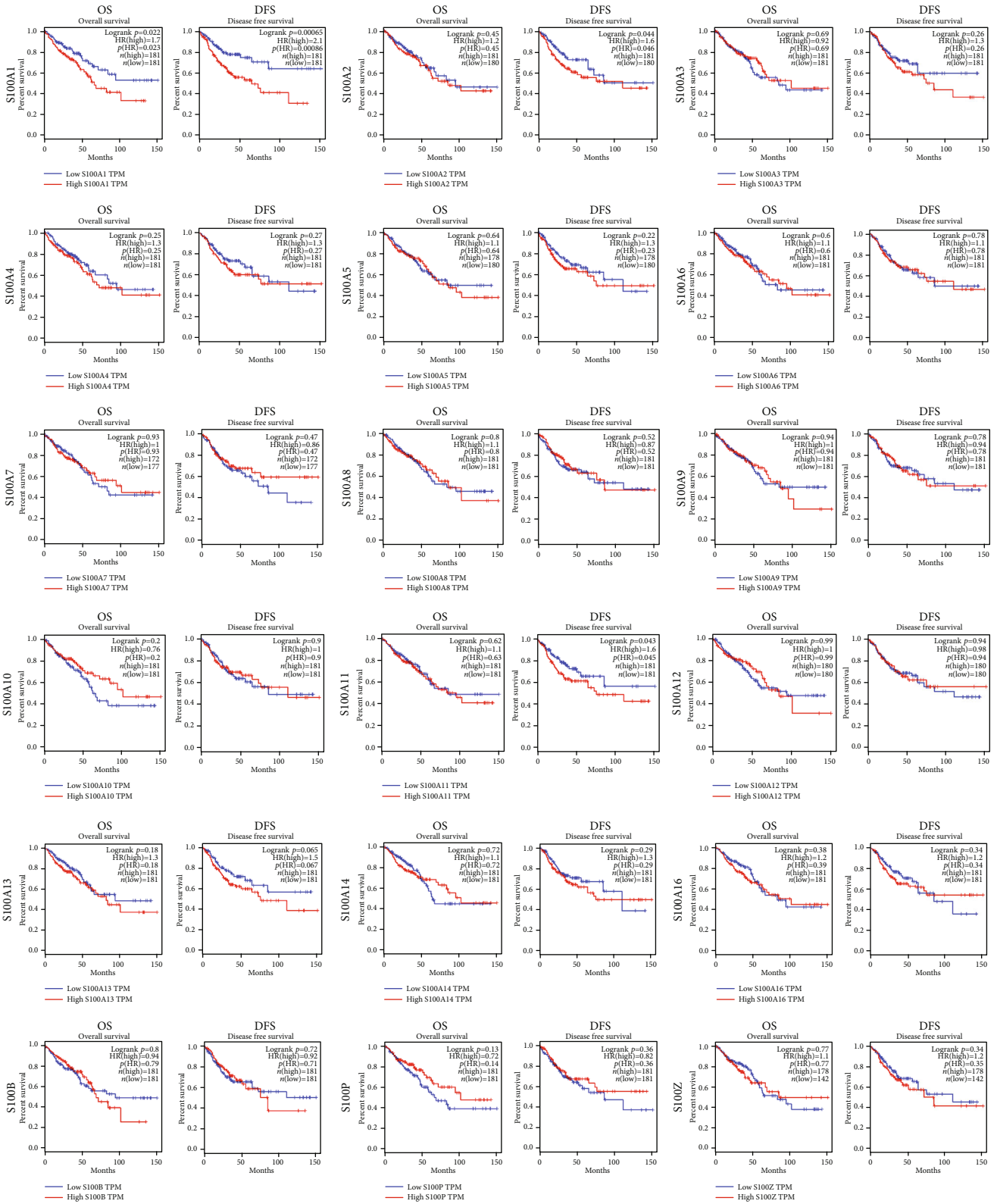


FIGURE 5: Survival analysis of S100 gene family in patients with CRC (GEPIA).

4. Discussion

Many reports have documented that S100 gene dysregulation is related to several cancers [3, 7, 8, 10, 21]. Although the

roles of S100 genes in tumorigenesis and prognosis of human cancers have been partly confirmed [2, 9, 10, 21], further extensive bioinformatics analyses of the S100 family in CRC have not yet been performed. This study for the first time

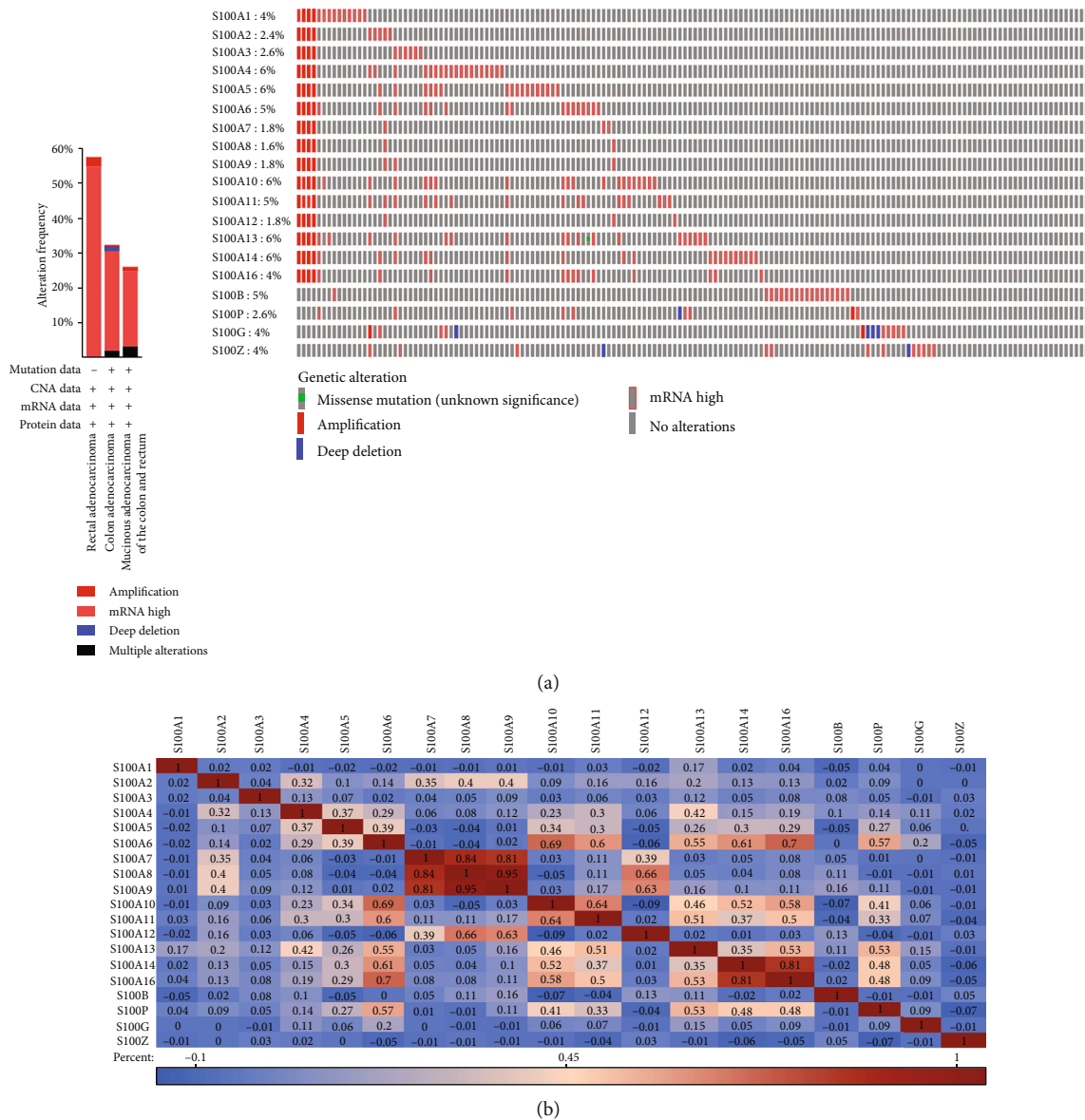


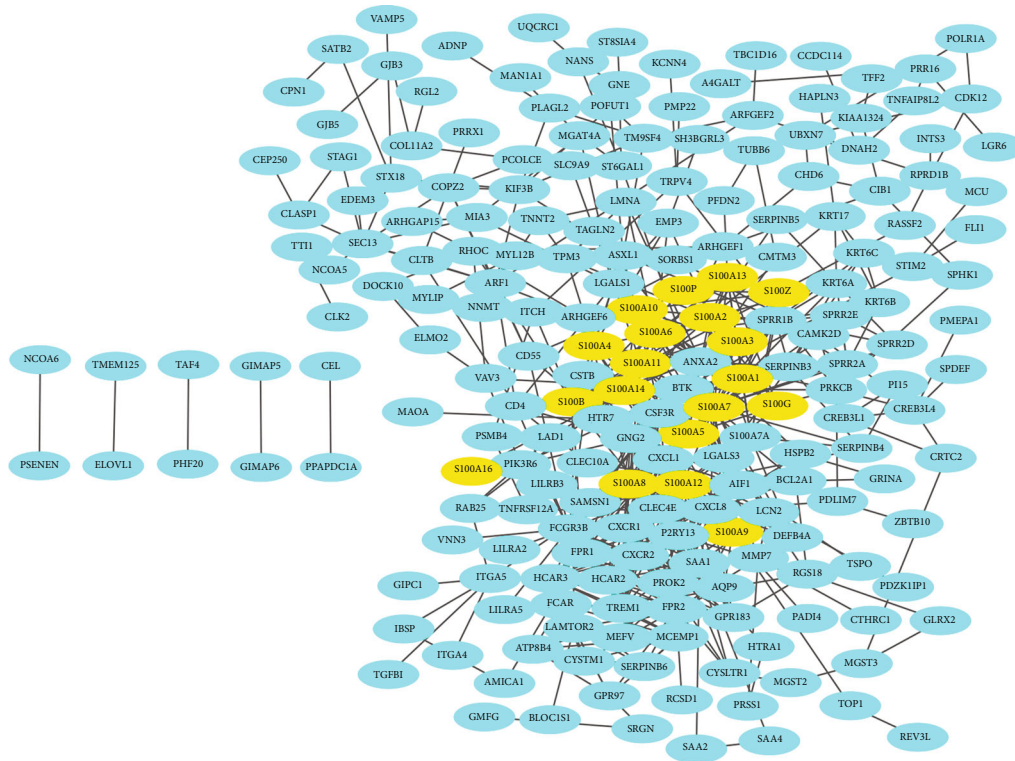
FIGURE 6: S100 gene expression and mutation analysis in COAD (cBioPortal). (a) Cancer type summary and oncoprint of the S100 members in cBioPortal. (b) Heatmap of the correlation among S100 members. Pearson correlation coefficients are exhibited as a color gradient from blue (negative correlation) to pink to reddish brown (positive correlation).

reports the prognostic (DFS, OS) values of the S100 gene family in CRC using bioinformatics tools, and our findings will underpin further studies on the mechanisms of dysregulated S100 genes in CRC, therapeutic targets, and optimization of treatment plans with improved prognosis. As analyses for S100A7A, S100A7L2, and S100G expressions in CRC are unavailable due to a lack of data, we merely focus on other members of the S100 family that are obviously related to the progression of CRC.

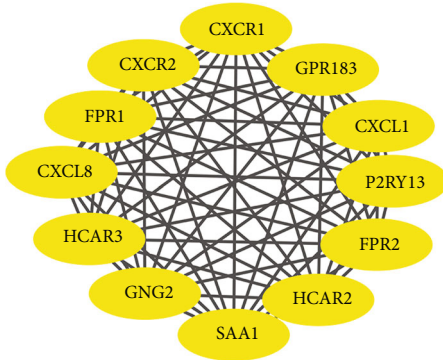
In our study, gene expression analyses show that S100A2, S100A11, and S100P expression levels in CRC tissues are significantly higher than those in noncancerous tissues, and S100A3 and S100A9 mRNAs are highly expressed in cancer tissues compared with normal tissue controls. By contrary, S100B is significantly downexpressed in CRC tissues. However, OncoPrint analysis and GEPIA

have yielded inconsistent results of S100A8, S100A10, S100A14, and S100A16 expression levels. Based on the OncoPrint database, there are no obvious distinctions in S100A1, S100A4, S100A5, S100A6, and S100A13 expressions between cancer tissues and normal colon mucosa, and S100A7, S100A12, and S100Z are slightly overexpressed in CRC tissues. Whereas, GEPIA shows that S100A5 and S100A6 are overexpressed in COAD and READ tissues, and S100A1, S100A4, S100A7, S100A12, S100A13, and S100Z expression levels are downregulated in cancerous tissues.

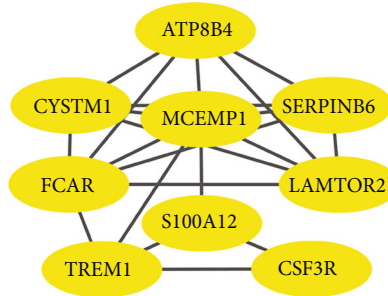
As for the prognostic value of dysregulated S100 genes, our results show a significant correlation between S100A10 and CRC at different stages of progression ($p = 0.0173$). However, such a strong correlation has not been observed in any other member of the S100 family. Notably, the



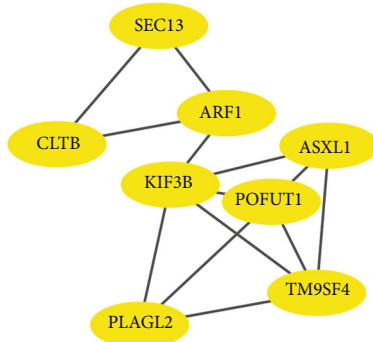
(a)



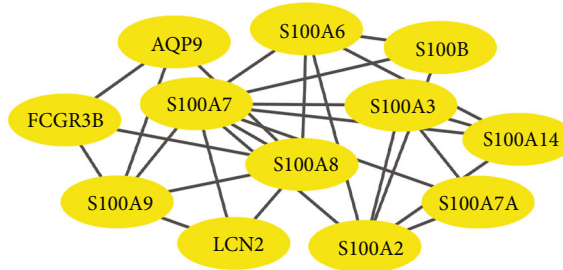
(b)



(c)



(d)



(e)

FIGURE 7: Continued.

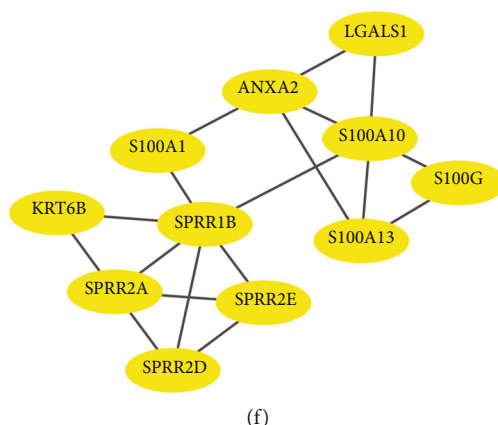


FIGURE 7: PPI network constructed by STRING and module composition visualized by Cytoscape. (a) Visualization the PPI network of S100 members and their mutation-related genes. S100 members are circled in yellow. (b) Module 1 of the top 5 modules with maximum depth of 100, node score cut-off of 0.2, k-core of 2, and degree cutoff of 2. (c) Module 2 of the top 5 modules. (d) Module 4 of the top 5 modules. (e) Module 3 of the top 5 modules. (f) Module 5 of the top 5 modules.

TABLE 2: Gene Ontology analysis of the S100 genes and their most significantly coexpressed genes in CRC.

Category	Term	Count	%	<i>p</i> value	FDR
GOTERM_BP_DIRECT	GO:0006954~inflammatory response	9	17.30769	9.11E-06	0.012361
GOTERM_BP_DIRECT	GO:0006935~chemotaxis	6	11.53846	2.24E-05	0.030396
GOTERM_BP_DIRECT	GO:0018149~peptide cross-linking	4	7.692308	0.000367	0.497246
GOTERM_BP_DIRECT	GO:0070098~chemokine-mediated signaling pathway	4	7.692308	0.001028	1.385837
GOTERM_BP_DIRECT	GO:0007186~G-protein coupled receptor signaling pathway	9	17.30769	0.003201	4.257847
GOTERM_BP_DIRECT	GO:0007204~positive regulation of cytosolic calcium ion concentration	4	7.692308	0.006236	8.138288
GOTERM_CC_DIRECT	GO:0070062~extracellular exosome	22	42.30769	5.91E-06	0.006523
GOTERM_CC_DIRECT	GO:0001533~cornified envelope	4	7.692308	0.000271	0.299178
GOTERM_CC_DIRECT	GO:0005886~plasma membrane	23	44.23077	0.000633	0.696457
GOTERM_CC_DIRECT	GO:0005576~extracellular region	12	23.07692	0.003667	3.973017
GOTERM_CC_DIRECT	GO:0048471~perinuclear region of cytoplasm	6	11.53846	0.027144	26.19097
GOTERM_CC_DIRECT	GO:0005887~integral component of plasma membrane	9	17.30769	0.037498	34.41178
GOTERM_MF_DIRECT	GO:0050786~RAGE receptor binding	7	13.46154	1.99E-13	2.24E-10
GOTERM_MF_DIRECT	GO:0005509~calcium ion binding	15	28.84615	9.3E-09	1.05E-05
GOTERM_MF_DIRECT	GO:0044548~S100 protein binding	4	7.692308	6.44E-06	0.007239
GOTERM_MF_DIRECT	GO:0048306~calcium-dependent protein binding	4	7.692308	0.000634	0.710203
GOTERM_MF_DIRECT	GO:0019959~interleukin-8 binding	2	3.846154	0.008683	9.337967
GOTERM_MF_DIRECT	GO:0035662~Toll-like receptor 4 binding	2	3.846154	0.011561	12.253

TABLE 3: KEGG pathway analysis of the S100 genes and their related coexpressed genes in colorectal cancer.

Category	Term	Count	%	<i>p</i> value	FDR
KEGG_PATHWAY	hsa04062: chemokine signaling pathway	5	9.615385	0.001459	1.441856
KEGG_PATHWAY	hsa04060: cytokine-cytokine receptor interaction	5	9.615385	0.003873	3.787008

Since the expression levels of the S100 genes are not completely parallel in the two databases, we mainly focused on prognostic S100 members that are consistent in gene expression levels. S100A1 proteins are abundantly expressed in the central neuronal system, heart muscle, and skeletal muscle [22]. Although S100A1 is proved to be a biomarker in human cancers, its role in colon cancers has been rarely been studied. S100A1 protein expressions are marginally higher in the colon connective tissues of normal samples and adenoma with low-grade dysplasia than CRC tissues and high-grade dysplastic lesions [23]. Bronckart et al. report the presence of S100A1 expression in node-negative colon cancer and S100A1 deficiency in node-positive colon cancer [23]. This indicates that S100A1 can be a candidate biomarker for the prognosis of early-stage colon cancer.

S100A2 gene expressions in colon cancers have also been reported [24] and are associated with poor OS and DFS of CRC patients [25, 26]. The high mRNA expression of S100A2 is associated with poor relapse-free survival, suggesting that S100A2 can be an independent risk factor for the recurrence of advanced CRC patients [27]. However, S100A2 as a predictor of stage progression in CRC has not been proven.

S100A3 plays an important role in tumorigenesis and progression of a variety of human cancers [28–30]. Activated and overexpressed S100A3 is associated with tumorigenesis, tumor occurrence, and progression of CRC [31], and S100A3 may be a potential target for CRC treatment. Consistently, our finding showed that the S100A3 overexpression predicted poor DFS of CRC patients ($p = 0.26$).

S100A8 and S100A9 which are mainly expressed in myeloid cells naturally form a stable heterodimer and involve in inflammatory processes that lead to autoimmune diseases and many human cancers [32, 33]. S100A8 and S100A9 have been proposed as crucial proinflammatory factors and contribute to premetastatic niche formation in CRC, which are consistent with our finding that S100A8 and S100A9 show inflammatory chemotactic effects in CRC. Kim et al. reveal that S100A8/9 heterocomplexes are upregulated in colon cancers and promote tumor progression [34]. However, the heterocomplex shows nonsignificant prognostic values in CRC in our study.

S100A10 intracellularly colocalizes with annexin A2 and involves in the translocation of S100A10 to the cytosolic face of the plasma membrane [35]. Zhang et al. report that S100A10 is correlated with cellular invasiveness, angiogenesis, and metastasis of CRC cells [31, 36]. Shang et al. find that S100A10 overexpressions in CRC can enhance oxaliplatin (L-OHP) sensitivity [37, 38], which is consistent with our results that S100A10 overexpressions significantly associate with longer OS of CRC patients. S100A11 is located in the cytoplasm of tumor cells and highly expressed in CRC tissues compared with adjacent normal tissues. This suggests that S100A11 involves in the cellular growth of progressive CRC [39, 40]. S100A13 is considered to be a potent angiogenic biomarker for astrocytic gliomas and melanoma, but its role in CRC is rarely reported [41, 42].

S100B alone can significantly increase proliferation and angiogenesis in intestinal colon cancer Caco-2 cells, which is considered to be an “ideal bridge” linking colonic inflammation and cancer [43]. Seguella et al. show that S100B markedly increases cell proliferation and invasiveness in CRC cells. Moreover, overexpressed S100B is implicated in postoperative relapse and a poor prognosis in CRC [44]. In our study, though S100B is significantly downexpressed in CRC tissues, contrary to our expectation, S100B suppression has no associations with stage progression, OS or DFS in CRC.

Emoto et al. first identified S100P as a new calcium-modulated protein in the human placenta in 2001 [1]. Previous evidences support that S100P protein and mRNA expressions in cancerous tissues significantly increase compared with normal colon mucosa tissues [45]. Wang et al. report that stage I-III CRC patients with positive S100P protein expressions exhibited shorter OS compared with negative S100P expressions. However, in our research, patients having higher S100P levels show an overall trend of better OS and DFS, without significant differences.

Besides, limitations in our study must be acknowledged. First, as differences between samples and data resources are inevitable, same genes that are inconsistently expressed in the two databases may result in cognitive confusion. Second, the gene expression analyses are performed based on online databases, which means our findings must be verified in more large-sample clinical trials on CRC.

5. Conclusion

In summary, we have systematically analyzed expressions of 21 genes in the S100 family and explored their prognostic value in CRC by using the OncoPrint and GEPIA databases, STRING, Cytoscape, cBioportal, and the DAVID database. Among the 21 S100 genes, 3 (S100A1, S100A2, and S100A11) are significantly associated with the prognosis of CRC patients, and only S100A10 is significantly correlated with CRC stage and progression, suggesting that S100A1, S100A2, and S100A11 can serve as potential prognostic markers. Therefore, the prognostic value of the S100 family, especially S100A10, needs to be verified in animal experiments and clinical trials. Our study will underpin researches on molecular mechanisms of S100 proteins and relevant signaling pathways in CRC progression.

Our research offers novel insights into the contribution of the S100 family to the prognosis and progression of CRC and paves a way for new S100-targeted therapies for CRC.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request. The data could also obtain from following open online websites, OncoPrint (www.oncoPrint.org), GEPIA, available at <http://gepia.cancer-pku.cn/> (March 11, 2020), cBioPortal (<http://www.cbioportal.org/>; accessed March 11, 2020), The Search Tool for the Retrieval of Interacting Genes/Proteins (STRING, <https://string-db.org/>, version 11.0; accessed

March 11, 2020) and The Database for Annotation, Visualization and Integrated Discovery (DAVID, <https://david.ncifcrf.gov/>; version 6.8; accessed March 11, 2020).

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Authors' Contributions

ML Z and J L conceived and designed the idea to this paper; PC S, Z L, and YL K participated in its design and coordination and supervised the study. ML Z and PC S collected and analyzed the data and drafted the paper; YP C and XJ Z analyzed the data and revised the final paper. All authors read and approved the final version of the manuscript. Meng-Lu Zeng was the first author of this article, Meng-Lu Zeng and Xian-Jin Zhu contributed equally to this work.

Acknowledgments

We would like to thank Yanfang Song, Clinical Laboratory, People's Hospital Affiliated to Fujian University of Traditional Chinese Medicine, for her valuable comments on the revision of the manuscript. The present study was supported by the Fujian Province Joint Funds for the Innovation of Science and Technology (grant no. 2017Y9051, 2019Y9060), the Fujian Medical Innovation Project (grant no. 2017-CX-20), and the Natural Science Foundation of Fujian Province (grant no. 2019J01151).

References

- [1] R. Donato, "S100: a multigenic family of calcium-modulated proteins of the ef-hand type with intracellular and extracellular functional roles," *The International Journal of Biochemistry & Cell Biology*, vol. 33, no. 7, pp. 637–668, 2001.
- [2] A. R. Bresnick, D. J. Weber, and D. B. Zimmer, "S100 proteins in cancer," *Nature Reviews. Cancer*, vol. 15, no. 2, pp. 96–109, 2015.
- [3] H. Chen, C. Xu, Q. Jin, and Z. Liu, "S100 protein family in human cancer," *American Journal of Cancer Research*, vol. 4, no. 2, pp. 89–115, 2014.
- [4] N. Ma, L. Zhu, L. Yang, Y. Cui, and Y. Zhan, "Prognostic values of s100 family mrna expression in ovarian cancer," *Cancer Biomarkers*, vol. 25, no. 1, pp. 67–78, 2019.
- [5] S. Zhang, Z. Wang, W. Liu et al., "Distinct prognostic values of s100 mrna expression in breast cancer," *Scientific Reports*, vol. 7, no. 1, 2017.
- [6] X. Wang, J. Han, D. B. Hardie, J. Yang, and C. H. Borchers, "The use of matrix coating assisted by an electric field (mcaef) to enhance mass spectrometric imaging of human prostate cancer biomarkers," *Journal of Mass Spectrometry*, vol. 51, no. 1, pp. 86–95, 2016.
- [7] P. Moravkova, D. Kohoutova, S. Rejchrt, J. Cyrany, and J. Bures, "Role of s100 proteins in colorectal carcinogenesis," *Gastroenterology Research and Practice*, vol. 2016, 7 pages, 2016.
- [8] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2017," *CA: a Cancer Journal for Clinicians*, vol. 67, no. 1, pp. 7–30, 2017.
- [9] J. J. Y. Sung, J. Y. W. Lau, K. L. Goh, and W. K. Leung, "Increasing incidence of colorectal cancer in asia: implications for screening," *The Lancet Oncology*, vol. 6, no. 11, pp. 871–876, 2005.
- [10] P. Jones, J. E. Cade, C. E. L. Evans, N. Hancock, and D. C. Greenwood, "Does adherence to the world cancer research fund/american institute of cancer research cancer prevention guidelines reduce risk of colorectal cancer in the uk women's cohort study?," *The British Journal of Nutrition*, vol. 119, no. 3, pp. 340–348, 2018.
- [11] Z. Tang, C. Li, B. Kang, G. Gao, C. Li, and Z. Zhang, "Gepia: a web server for cancer and normal gene expression profiling and interactive analyses," *Nucleic Acids Research*, vol. 45, no. W1, pp. W98–w102, 2017.
- [12] K. Tomczak, P. Czerwińska, and M. Wiznerowicz, "The cancer genome atlas (tcga): an immeasurable source of knowledge," *Contemporary oncology*, vol. 19, no. 1a, pp. A68–A77, 2015.
- [13] D. Szklarczyk, A. Franceschini, S. Wyder et al., "String v10: protein-protein interaction networks, integrated over the tree of life," *Nucleic Acids Research*, vol. 43, no. D1, pp. D447–D452, 2015.
- [14] M. Ashburner, C. A. Ball, J. A. Blake et al., "Gene ontology: tool for the unification of biology," *Nature Genetics*, vol. 25, no. 1, pp. 25–29, 2000.
- [15] M. Kanehisa and S. Goto, "Kegg: kyoto encyclopedia of genes and genomes," *Nucleic Acids Research*, vol. 28, no. 1, pp. 27–30, 2000.
- [16] S. Kaiser, Y. K. Park, J. L. Franklin et al., "Transcriptional recapitulation and subversion of embryonic colon development by mouse colon tumor models and human colon cancer," *Genome Biology*, vol. 8, no. 7, p. R131, 2007.
- [17] J. Gaedcke, M. Grade, K. Jung et al., "Mutated kras results in overexpression of dusp4, a map-kinase phosphatase, and smyd3, a histone methyltransferase, in rectal carcinomas," *Genes, Chromosomes & Cancer*, vol. 49, no. 11, pp. 1024–1034, 2010.
- [18] M. Skrzypczak, K. Goryca, T. Rubel et al., "Modeling oncogenic signaling in colon tumors by multidirectional analyses of microarray data directed for maximization of analytical reliability," *PLoS One*, vol. 5, no. 10, p. e13091, 2010.
- [19] D. H. Ki, H. C. Jeung, C. H. Park et al., "Whole genome analysis for liver metastasis gene signatures in colorectal cancer," *International Journal of Cancer*, vol. 121, no. 9, pp. 2005–2012, 2007.
- [20] D. A. Notterman, U. Alon, A. J. Sierk, and A. J. Levine, "Transcriptional gene expression profiles of colorectal adenoma, adenocarcinoma, and normal tissue examined by oligonucleotide arrays," *Cancer Research*, vol. 61, no. 7, pp. 3124–3130, 2001.
- [21] S. R. Gross, C. G. T. Sin, R. Barraclough, and P. S. Rudland, "Joining s100 proteins and migration: for better or for worse, in sickness and in health," *Cellular and Molecular Life Sciences*, vol. 71, no. 9, pp. 1551–1579, 2014.
- [22] N. T. Wright, B. R. Cannon, D. B. Zimmer, and D. J. Weber, "S100a1: structure, function, and therapeutic potential," *Current Chemical Biology*, vol. 3, no. 2, pp. 138–145, 2009.
- [23] Y. Bronckart, C. Decaestecker, N. Nagy et al., "Development and progression of malignancy in human colon tissues are

- correlated with expression of specific ca(2+)-binding s100 proteins," *Histology and Histopathology*, vol. 16, no. 3, pp. 707–712, 2001.
- [24] M. D. Giráldez, J. J. Lozano, M. Cuatrecasas et al., "Gene-expression signature of tumor recurrence in patients with stage ii and iii colon cancer treated with 5'fluoruracil-based adjuvant chemotherapy," *International Journal of Cancer*, vol. 132, no. 5, pp. 1090–1097, 2013.
- [25] N. M. Alajez, "Large-scale analysis of gene expression data reveals a novel gene expression signature associated with colorectal cancer distant recurrence," *PLoS One*, vol. 11, no. 12, article e0167455, 2016.
- [26] N. Long, S. Park, N. Anh et al., "High-throughput omics and statistical learning integration for the discovery and validation of novel diagnostic signatures in colorectal cancer," *International Journal of Molecular Sciences*, vol. 20, no. 2, p. 296, 2019.
- [27] A. Masuda, T. Ishikawa, K. Mogushi et al., "Overexpression of the s100a2 protein as a prognostic marker for patients with stage ii and iii colorectal cancer," *International Journal of Oncology*, vol. 48, no. 3, pp. 975–982, 2016.
- [28] R. Tao, Z. F. Wang, W. Qiu et al., "Role of s100a3 in human hepatocellular carcinoma and the anticancer effect of sodium cantharidinate," *Experimental and Therapeutic Medicine*, vol. 13, no. 6, pp. 2812–2818, 2017.
- [29] B. Liu, W.-Y. Sun, C.-Y. Zhi et al., "Role of s100a3 in human colorectal cancer and the anticancer effect of cantharidinate," *Experimental and Therapeutic Medicine*, vol. 6, no. 6, pp. 1499–1503, 2013.
- [30] M. Kang, H. S. Lee, Y. J. Lee et al., "S100a3 suppression inhibits in vitro and in vivo tumor growth and invasion of human castration-resistant prostate cancer cells," *Urology*, vol. 85, no. 1, pp. 273.e9–273.e15, 2015.
- [31] L. Zhang, D. K. Fogg, and D. M. Waisman, "Rna interference-mediated silencing of the s100a10 gene attenuates plasmin generation and invasiveness of colo 222 colorectal cancer cells," *The Journal of Biological Chemistry*, vol. 279, no. 3, pp. 2053–2062, 2004.
- [32] S. Y. Lim, A. E. Yuzhalin, A. N. Gordon-Weeks, and R. J. Muschel, "Tumor-infiltrating monocytes/macrophages promote tumor invasion and migration by upregulating s100a8 and s100a9 expression in cancer cells," *Oncogene*, vol. 35, no. 44, pp. 5735–5745, 2016.
- [33] C. Gebhardt, J. Németh, P. Angel, and J. Hess, "S100a8 and s100a9 in inflammation and cancer," *Biochemical Pharmacology*, vol. 72, no. 11, pp. 1622–1631, 2006.
- [34] H. J. Kim, H. J. Kang, H. Lee et al., "Identification of s100a8 and s100a9 as serological markers for colorectal cancer," *Journal of Proteome Research*, vol. 8, no. 3, pp. 1368–1379, 2009.
- [35] M. Kwon, C. S. Yoon, W. Jeong, S. G. Rhee, and D. M. Waisman, "Annexin a2-s100a10 heterotetramer, a novel substrate of thioredoxin," *The Journal of Biological Chemistry*, vol. 280, no. 25, pp. 23584–23592, 2005.
- [36] K. Mijung, T. J. MacLeod, Y. Zhang, and D. M. Waisman, "S100a10, annexin a2, and annexin a2 heterotetramer as candidate plasminogen receptors," *Frontiers in Bioscience*, vol. 10, no. 1-3, p. 300, 2005.
- [37] J. Shang, Z. Zhang, W. Song et al., "S100a10 as a novel biomarker in colorectal cancer," *Tumour Biology*, vol. 34, no. 6, pp. 3785–3790, 2013.
- [38] S. Suzuki and Y. Tanigawara, "Forced expression of s100a10 reduces sensitivity to oxaliplatin in colorectal cancer cells," *Proteome Science*, vol. 12, no. 1, p. 26, 2014.
- [39] H. He, J. Li, S. Weng, M. Li, and Y. Yu, "S100a11: diverse function and pathology corresponding to different target proteins," *Cell Biochemistry and Biophysics*, vol. 55, no. 3, pp. 117–126, 2009.
- [40] G. Wang, X. Wang, S. Wang et al., "Colorectal cancer progression correlates with upregulation of s100a11 expression in tumor tissues," *International Journal of Colorectal Disease*, vol. 23, no. 7, pp. 675–682, 2008.
- [41] D. Massi, M. Landriscina, A. Piscazzi et al., "S100a13 is a new angiogenic marker in human melanoma," *Modern Pathology*, vol. 23, no. 6, pp. 804–813, 2010.
- [42] M. Landriscina, G. Schinzari, G. Di Leonardo et al., "S100a13, a new marker of angiogenesis in human astrocytic gliomas," *Journal of Neuro-Oncology*, vol. 80, no. 3, pp. 251–259, 2006.
- [43] L. Seguella, R. Capuano, M. Pesce et al., "S100b protein stimulates proliferation and angiogenic mediators release through raga/pakt/mtor pathway in human colon adenocarcinoma caco-2 cells," *International Journal of Molecular Sciences*, vol. 20, no. 13, p. 3240, 2019.
- [44] M. Y. Huang, H. M. Wang, T. S. Tok et al., "EVI2B, ATP2A2, S100B, TM4SF3, and OLFM4As potential prognostic markers for postoperative taiwanese colorectal cancer patients," *DNA and Cell Biology*, vol. 31, no. 4, pp. 625–635, 2012.
- [45] L. Dong, F. Wang, X. Yin et al., "Overexpression of s100p promotes colorectal cancer metastasis and decreases chemosensitivity to 5-fu in vitro," *Molecular and Cellular Biochemistry*, vol. 389, no. 1-2, pp. 257–264, 2014.