



Weighted Gene Co-expression Network Analysis Identifies a Cancer-Associated Fibroblast Signature for Predicting Prognosis and Therapeutic Responses in Gastric Cancer

Hang Zheng^{1†}, Heshu Liu^{2†}, Huayu Li¹, Weidong Dou¹ and Xin Wang^{1*}

OPEN ACCESS

Edited by:

José Alexandre Ferreira,
Portuguese Oncology Institute,
Portugal

Reviewed by:

Manish Charan,
The Ohio State University,
United States
Ila Pant,
Icahn School of Medicine at Mount
Sinai, United States

*Correspondence:

Xin Wang
wangxin_guo@126.com

[†]These authors have contributed
equally to this work and share first
authorship

Specialty section:

This article was submitted to
Molecular Diagnostics and
Therapeutics,
a section of the journal
Frontiers in Molecular Biosciences

Received: 20 July 2021

Accepted: 01 September 2021

Published: 08 October 2021

Citation:

Zheng H, Liu H, Li H, Dou W and
Wang X (2021) Weighted Gene Co-
expression Network Analysis Identifies
a Cancer-Associated Fibroblast
Signature for Predicting Prognosis and
Therapeutic Responses in
Gastric Cancer.
Front. Mol. Biosci. 8:744677.
doi: 10.3389/fmolb.2021.744677

¹Department of General Surgery, Peking University First Hospital, Peking University, Beijing, China, ²Department of Oncology, Beijing Chaoyang Hospital, Capital Medical University, Beijing, China

Background: Cancer-associated fibroblasts (CAFs) are the most prominent cellular components in gastric cancer (GC) stroma that contribute to GC progression, treatment resistance, and immunosuppression. This study aimed at exploring stromal CAF-related factors and developing a CAF-related classifier for predicting prognosis and therapeutic effects in GC.

Methods: We downloaded mRNA expression and clinical information of 431 GC samples from Gene Expression Omnibus (GEO) and 330 GC samples from The Cancer Genome Atlas (TCGA) databases. CAF infiltrations were quantified by the estimate the proportion of immune and cancer cells (EPIC) method, and stromal scores were calculated via the Estimation of STromal and Immune cells in MAlignant Tumors using Expression data (ESTIMATE) algorithm. Stromal CAF-related genes were identified by weighted gene co-expression network analysis (WGCNA). A CAF risk signature was then developed using the univariate and least absolute shrinkage and selection operator method (LASSO) Cox regression model. We applied the Spearman test to determine the correlation among CAF risk score, CAF markers, and CAF infiltrations (estimated *via* EPIC, xCell, microenvironment cell populations-counter (MCP-counter), and Tumor Immune Dysfunction and Exclusion (TIDE) algorithms). The TIDE algorithm was further used to assess immunotherapy response. Gene set enrichment analysis (GSEA) was applied to clarify the molecular mechanisms.

Results: The 4-gene (COL8A1, SPOCK1, AEBP1, and TIMP2) prognostic CAF model was constructed. GC patients were classified into high- and low-CAF-risk groups in accordance with their median CAF risk score, and patients in the high-CAF-risk group had significant worse prognosis. Spearman correlation analyses revealed the CAF risk score was strongly and positively correlated with stromal and CAF infiltrations, and the four model genes also exhibited positive correlations with CAF markers. Furthermore, TIDE analysis revealed high-CAF-risk patients were less likely to respond to immunotherapy.

GSEA revealed that epithelial–mesenchymal transition (EMT), TGF- β signaling, hypoxia, and angiogenesis gene sets were significantly enriched in high–CAF-risk group patients.

Conclusion: The present four-gene prognostic CAF signature was not only reliable for predicting prognosis but also competent to estimate clinical immunotherapy response for GC patients, which might provide significant clinical implications for guiding tailored anti-CAF therapy in combination with immunotherapy for GC patients.

Keywords: gastric cancer, cancer-associated fibroblasts, weighted gene co-expression network analysis, biomarker, prognosis, immunotherapy

INTRODUCTION

Gastric cancer (GC) ranks fifth among the most common cancers and is the fourth leading cause of cancer-related mortality worldwide (Sung et al., 2021). Leaving aside improvements in gastroscopic screening and various treatment strategies, recurrence and metastasis remain the main causes of GC death, and the current therapeutic efficacy on recurrent and metastatic GC is still unsatisfactory (Lee et al., 2016; Thrift and El-Serag, 2020). GC tissues are composed of neoplastic cancer cells as well as the immune and stromal milieu where tumor cells are located, which is termed as tumor microenvironment (TME). Accumulating evidence indicated tumor stromal components in TME are critical for tumor growth and metastasis, immunosuppression, and drug resistance (Hanahan and Coussens, 2012; Quail and Joyce, 2013), which have embraced a spacious field of investigation.

As the most prominent cell type of tumor stroma, cancer-associated fibroblasts (CAFs) are crucial sources of growth factors and cytokines that promote tumor progression and migration (Kojima et al., 2010; Tommelein et al., 2015), stimulate epithelial–mesenchymal transition (EMT) (Wu et al., 2017; Fiori et al., 2019), and induce chemoresistance (Lotti et al., 2013; Li et al., 2016) and immunosuppression (Kraman et al., 2010; Monteran and Erez, 2019). CAFs are also capable of depositing and reorganizing the extracellular matrix (ECM), which serves as a thick physical barrier that supports tumor cell invasion and restrains the infiltrations of antitumor leukocytes, leading to tumor progression, immune evasion, and therapy resistance (Ma et al., 2016; Lakins et al., 2018; Kaur et al., 2019; Gamradt et al., 2021). Thus, targeting CAF-mediated immunosuppressive stromal microenvironment in combination with immunotherapy could promisingly ameliorate the response to immune checkpoint inhibitors. For instance, exhaustion of fibroblast activation protein (FAP)–positive CAFs in murine models led to increased CD8⁺ T cell infiltrations and decreased macrophages proportions (Duperret et al., 2018), and therapeutic effects of anti-CTLA4 and anti-PD-1 were consequently enhanced (Feig et al., 2013). Unfortunately, FAP-based drugs against CAFs failed to pass Phase II trials owing to the unsatisfactory clinical response in metastatic colorectal cancer patients (Hofheinz et al., 2003; Narra et al., 2007), and such a CAF inhibitory strategy is currently lacking in GC treatment. In this regard, it is imperative to explore stromal CAF-related factors in GC.

Weighted gene co-expression network analysis (WGCNA) is a systematic bioinformatics algorithm that is competent to incorporate highly and coordinately expressed genes into several gene modules and investigate the module's relationships with the phenotype of interest (Langfelder and Horvath, 2008). WGCNA has been successfully applied for identifying CAF markers (Liu et al., 2021a; Liu et al., 2021b). So far, CAF and stromal infiltrations have not been subjected to WGCNA analysis in GC. In this study, for the first time, WGCNA was employed simultaneously on two transcriptome datasets collected from publicly available Gene Expression Omnibus (GEO) and The Cancer Genome Atlas (TCGA) databases. We detected hub modules that were most correlated with stromal CAF infiltrations. Then, by applying univariate and Least Absolute Shrinkage and Selection Operator (LASSO) Cox regression analyses, we identified COL8A1, SPOCK1, AEBP1, and TIMP2 as prognostic CAF markers and constructed the four-gene CAF signature capable of predicting prognosis and therapeutic responses in GC. Our results hint that the CAF model might be a novel anti-CAF therapeutic approach in GC.

MATERIALS AND METHODS

Data Acquisition and Preprocessing

The fragments per kilobase of transcript per million mapped reads (FPKM) format RNA-seq data and corresponding prognostic data (follow-up time more than 30 days) of 330 TCGA stomach adenocarcinoma (TCGA-STAD) samples were downloaded through UCSC Xena browser (GDC hub) (<https://gdc.xenahubs.net>) (Goldman et al., 2020). The normalized FPKM values were converted to transcripts per million (TPM) and $\log_2(\text{TPM}+1)$ transformed (Wagner et al., 2012). We also obtained normalized expression data and clinical information of 431 GC samples in GSE84437 from the GEO database (Yoon et al., 2020). The highest value was reserved if one gene matched multiple probes.

CAF Infiltration Estimation and Stromal Score Calculation

CAF abundances were separately estimated *via* four methods: cell-type deconvolution (constrained least square optimization)–based Estimate the Proportion of Immune and Cancer cells (EPIC) algorithm (Racle et al., 2017), gene signature enrichment–based xCell algorithm (Aran et al., 2017), marker

genes expressions-based microenvironment cell populations-counter (MCP-counter) (Becht et al., 2016), and Tumor Immune Dysfunction and Exclusion (TIDE) algorithms (Jiang et al., 2018). The first three methods were achieved *via* a `deconvolve()` function of `immunedeconv` R package (version 2.0.3) (Sturm et al., 2020), and the TIDE method was implemented through <http://tide.dfc.harvard.edu/>. In addition, the Estimation of STromal and Immune cells in Malignant Tumor tissues using Expression data (ESTIMATE) algorithm was applied to calculate the stromal score *via* `estimate` R package (version 1.0.13), which indicates the stromal infiltrating levels of each tumor sample (Yoshihara et al., 2013).

CAF and Stromal Co-expression Network Constructions

Co-expression networks and hub genes that targeted CAF infiltrations as well as stromal scores were constructed and detected *via* WGCNA R package (version 1.68) (Langfelder and Horvath, 2008). Genes with the top 5,000 of median absolute deviation (MAD) were first chosen as the input genes for network constructions in both TCGA-STAD and GSE84437 cohorts. Then, the Pearson's correlation similarity matrix between any gene pairs was calculated (s_{ij} , where ij represents pair-wise genes) and increased to soft-thresholding power β (s_{ij}^β) based on the scale-free topology network criterion. Subsequently, the adjacency matrix was clustered using topological overlap measure (TOM) and dissimilarity (1-TOM) between genes, and we conducted a dynamic tree cut algorithm on the dendrogram for gene module identifications with minimum gene numbers as 30 in each module. Each module expression's first principal component was summarized as module eigengenes (MEs), the Pearson's correlations between MEs and EPIC-quantified CAF infiltrations as well as the stromal score were evaluated, and the most correlated module was picked for further analysis. Then, we measured gene significance (GS) for the traits and module membership (MM indicates the correlation between ME and gene expression) of individual genes in the identified hub module, and hub genes were filtered out under the strict criteria of $GS > 0.4$ and $MM > 0.8$. Finally, the overlapping hub genes between TCGA-STAD and GSE84437 cohorts constituted the final hub genes.

Gene Ontology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) Analyses

GO and KEGG pathway enrichment analyses were performed on the final hub genes to identify the biological functions (including biological processes (BPs), molecular functions (MFs), and cellular components (CCs)) and pathways through `clusterProfiler` R package (version 3.14.3) (Yu et al., 2012). $p < 0.05$ was considered statistically enriched.

Prognostic Model Construction and Validation

The GSE84437 cohort was selected for CAF risk model construction owing to its larger sample size, while 330 cases

from TCGA-STAD were assigned to the validation cohort. The univariate Cox regression model was performed to identify prognostic stromal CAF hub genes on overall survival (OS); genes with $p < 0.05$ were subsequently put into LASSO Cox regression analysis with 1,000 iterations for gene reduction *via* `glmnet` R package (Simon et al., 2011). Then, the CAF risk model was constructed as follows: CAF risk score = $\sum (\beta_i * \text{Exp}_i)$, where β_i refers to the LASSO coefficient of i th gene, and Exp_i represents the i th gene's expression value. GC patients were classified into high- and low-CAF-risk groups based on their median CAF risk scores, and the OS difference between two groups was estimated *via* Kaplan-Meier curves and the log-rank test. Similarly, the CAF risk model was validated in the TCGA-STAD cohort.

CAF Markers Collections and Correlation Analysis

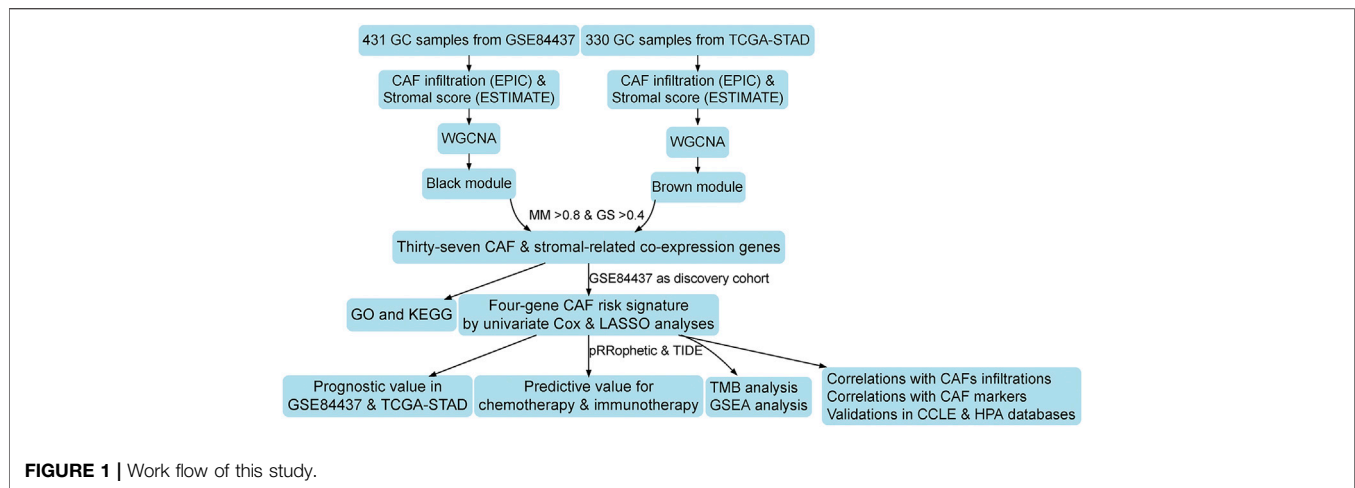
CAF specific and nonspecific markers were collected from published literature (Gascard and Tlsty, 2016; Han et al., 2020). To ensure the reliability of our CAF model markers in GC, we analyzed the Spearman's correlations between the CAF risk score and stromal score as well as multi-estimated CAF infiltrations (EPIC, xCell, MCP-counter, and TIDE). Correlations between CAF model genes and published CAF markers were also analyzed on both TCGA-STAD and GSE84437 cohorts.

Chemotherapy and Immunotherapy Response Predictions

Based on the largest publicly attainable pharmacogenomics database, Genomics of Drug Sensitivity in Cancer (GDSC) (<https://www.cancerrxgene.org/>) (Yang et al., 2013), half-maximal inhibitory concentration (IC50) values of common drugs (bleomycin, lapatinib, paclitaxel, camptothecin, cisplatin, docetaxel, methotrexate, and sunitinib) in each GC sample were estimated based on the transcriptome data by ridge regression with ten-fold cross-validation in `pRRophetic` R package (version 0.5) (Geeleher et al., 2014a; Geeleher et al., 2014b). Subsequently, the TIDE (<http://tide.dfc.harvard.edu/>) online algorithm was adopted for immune checkpoint blockade therapy response predictions (Jiang et al., 2018). Differences in response rates between high- and low-CAF-risk groups were examined by the chi-squared test, and the predictive efficacy of the CAF risk signature was evaluated by ROC curves and area under the curve (AUC) values.

Somatic Alteration Data Collection and Analyses

The somatic mutation data of the TCGA-STAD cohort were downloaded *via* the `GDCquery_Maf()` function (pipelines = "mutect2" (Cibulskis et al., 2013)) of `TCGAbiolinks` R package (Colaprico et al., 2016). The top 20 highest mutational frequencies in both low- and high-CAF-risk groups were recognized and visualized *via* `maftools` R package (Mayakonda et al., 2018). Tumor mutation burden (TMB) has been proposed as an immunotherapy efficacy predictor (Yarchoan et al., 2017),



and the TMB value of each STAD sample was then calculated *via* the `tmb()` function of `maftools` package, and Spearman's correlation between TMB and CAF risk scores were analyzed.

Enrichment Analyses

Gene set enrichment analysis (GSEA) was performed to explore the enriched hallmark and KEGG pathway gene sets between high- and low-CAF-risk groups in GSE84437 *via* `enrichplot` and `clusterProfiler` R packages. The “c2. cp.kegg.v7.4. symbols” and “h.all.v7.4. symbols” gene sets were derived from the Molecular Signatures Database (MSigDB) (Liberzon et al., 2015). Furthermore, ssGSEA was applied to calculate the enrichment scores of EMT, TGF- β , and angiogenesis hallmark gene sets (Hänzelmann et al., 2013). Spearman's correlations analysis was performed to assess the correlation between the CAF risk score and gene set enrichment scores.

Validation *via* Cancer Cell Line Encyclopedia (CCLE) and Human Protein Atlas (HPA) Databases

For cellular level validation, the mRNA expressions of the identified markers in 38 fibroblasts and 39 GC cell lines were downloaded from the CCLE database (<https://portals.broadinstitute.org/ccle>) (Ghandi et al., 2019), and we examined their expression patterns in fibroblasts and CRC cell lines *via* heat map and Wilcoxon tests. In addition, with respect to protein level investigation, immunohistochemical (IHC) staining images of these markers in GC tissues were downloaded from the HPA online database (<https://www.proteinatlas.org/>) (Uhlén et al., 2015), and the target protein localization could be directly observed.

Statistical Analysis

All statistical analyses were performed using R software (version 3.6.3; <https://www.r-project.org/>). The median CAF risk score was the cutoff value for each cohort in dividing GC patients into high- and low-CAF-risk subgroups. The Wilcoxon test was applied for pairwise comparisons. The Kaplan–Meier curve with the log-rank test was adopted for overall survival

comparisons *via* survival and `survminer` R packages. $p < 0.05$ was regarded as statistically significant.

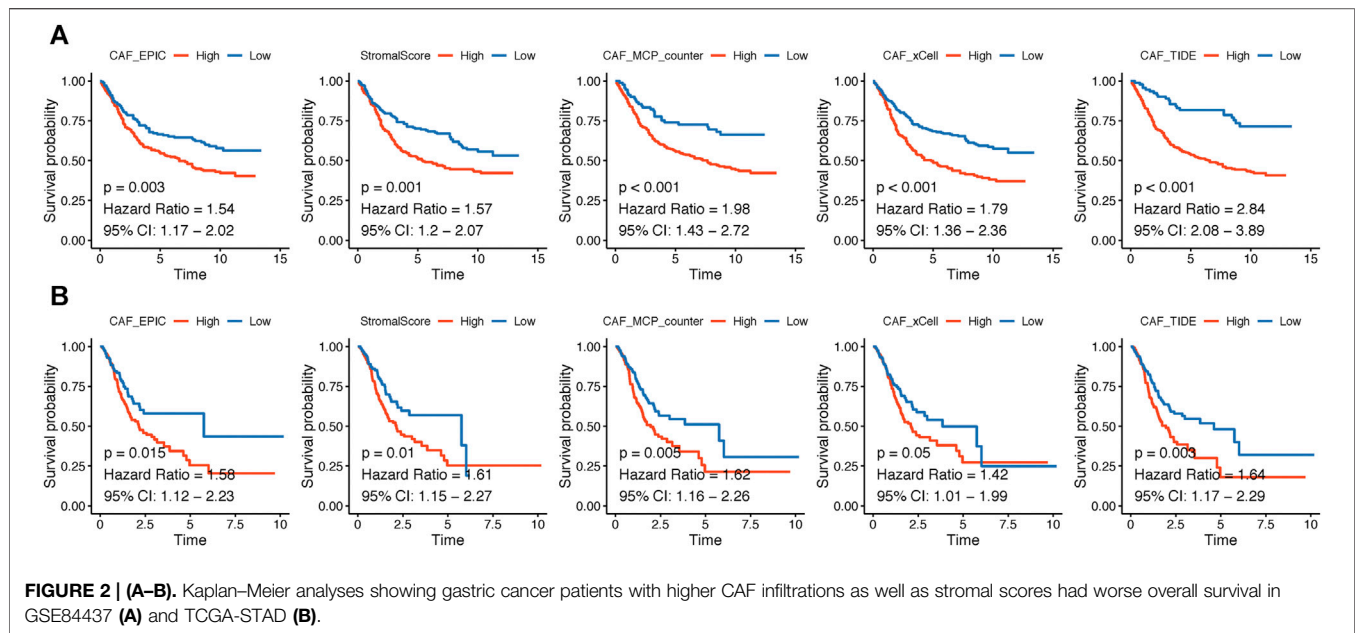
RESULTS

Higher CAF Infiltrations and Stromal Scores Indicate Worse OS in GC Patients

The flowchart of this research is displayed in **Figure 1**. CAF infiltrations were multiply predicted by EPIC, `xCell`, `MCP-counter`, and `TIDE` methods, and the stromal score was calculated by the `estimate` algorithm. Their prognostic values on OS were evaluated *via* log-rank tests; Kaplan–Meier curves indicated that higher CAF infiltrations and stromal scores were notably correlated with poorer OS of GC patients in both GSE84437 (**Figure 2A**) and TCGA-STAD (**Figure 2B**) cohorts, which highlighted the importance of further exploration of CAF and stromal-related genes for GC. Herein, the EPIC-estimated CAF abundances and stromal scores were summarized as phenotype data for the subsequent analysis, and the other three estimated CAF infiltrations data were utilized for the identified CAF model external validations.

Co-Expression Network of CAF and Stromal Scores

WGCNA analysis was performed in both GSE84437 and TCGA-STAD. To construct a scale-free topology network, the soft threshold power (β) of 8 in GSE84437 (scale-free $R^2 = 0.97$) (**Figure 3A**) and 6 in TCGA-STAD (scale-free $R^2 = 0.87$) (**Figure 3B**) was estimated. For GSE84437, the hierarchical clustering tree revealed that 11 co-expression models were clustered (**Figure 3C**), and the black module had the strongest positive correlation with the CAF proportion (Cor = 0.91, $P = 6e-161$) and stromal score (Cor = 0.84, $P = 3e-116$) (**Figure 3E**). For TCGA-STAD, the dynamic hybrid cutting clustered 9 co-expression models (**Figure 3D**), with the brown module having the strongest positive correlation with the CAF proportion (Cor = 0.88, $P = 4e-108$) and stromal score (Cor = 0.88, $P = 3e-104$) (**Figure 3F**). Hence, the two modules were focused for in-depth



investigations. A total of 356 and 302 genes were incorporated in the black and brown modules, respectively. In the black module, the scatter plots illustrated the strong correlations between MM and GS for CAF (Cor = 0.94, $p = 2.1 \times 10^{-167}$) and stromal scores (Cor = 0.85, $p = 1.4 \times 10^{-100}$) (Figure 3G); in the brown module, the strong correlations were also observed between MM and GS for CAF (Cor = 0.93, $P = 2 \times 10^{-132}$) and stromal scores (Cor = 0.95, $p = 1.1 \times 10^{-153}$) (Figure 3H). Then, by taking $MM > 0.8$ and $GS > 0.4$ as the threshold criteria, a total of 48 genes in the black model of GSE84437 and 101 genes in the brown module of TCGA-STAD were, respectively, screened out as hub genes which are highly correlated with CAF and stromal scores.

Functional Analyses of Hub Genes

As shown in Figure 4A, by the intersection of two hub gene sets, 37 common genes were detected and visualized via a Venn diagram. Subsequently, we performed GO and KEGG analyses on these 37 genes. Extracellular matrix organization and extracellular structure organization were the major enriched BP terms; the collagen-containing extracellular matrix and extracellular matrix structural constituents were the major enriched CC and MF terms, respectively. Protein digestion and absorption, focal adhesion, and the PI3K-Akt signaling pathway were the main enriched KEGG pathways.

Construction of the Stromal CAF-Based Prognostic Risk Model

Four hundred and thirty-one GC samples from GSE84437 were used as the training cohort owing to the larger sample size, and 330 TCGA-STAD samples were used as the validation group. By performing univariate Cox regression analysis of the 37 common hub genes, 33 OS-related genes with $p < 0.05$ were screened out and subjected to the following LASSO Cox regression analysis (Figures 4D,E). Four genes were finally identified for the CAF risk model

construction: CAF risk score = expression of COL8A1 * 0.1 + expression of SPOCK1 * 0.007 + expression of AEBP1 * 0.021 + expression of TIMP2 * 0.064 (Figure 4F). GC patients in each cohort were divided into high- and low-CAF-risk groups with the median risk score as the cutoff value. Kaplan–Meier curves revealed that GC patients in the high-CAF-risk group experienced worse OS than those in the low-CAF-risk group in both GSE84437 (HR = 1.768, 95%CI: 1.339–2.335, log-rank $p < 0.001$) (Figure 4G) and TCGA-STAD (HR = 1.522, 95%CI: 1.086–2.134, log-rank $p = 0.015$) (Figure 4H). These results indicated CAF and stromal-related signature genes were crucial prognostic markers in GC.

CAF Signature Genes Were Highly Correlated With CAF Infiltrations and CAF Markers

To further verify the robustness of the CAF model as an indicator in predicting CAF infiltrations, we performed Spearman's correlation analyses between the CAF risk score and stromal score as well as CAF abundances predicted by EPIC and other three methods: xCell, MCP-counter, and TIDE. Consistently, we observed the CAF risk score was strongly and positively correlated with multi-estimated CAF infiltrations and the stromal score in both GSE84437 (Figure 5A) and TCGA-STAD (Figure 5B) cohorts. Moreover, we observed the CAF risk score and the expression levels of the four genes were highly and positively correlated with a bunch of the collected CAF markers in both GSE84437 (Figures 5C,E) and TCGA-STAD (Figures 5D,F) cohorts.

Sensitivity of Chemotherapy and Immunotherapy Between CAF-Risk Groups

Adjuvant chemotherapy following radical surgery has been the standard approach regarding GC. IC50 values of several drugs mentioned in the methods section were estimated based on the

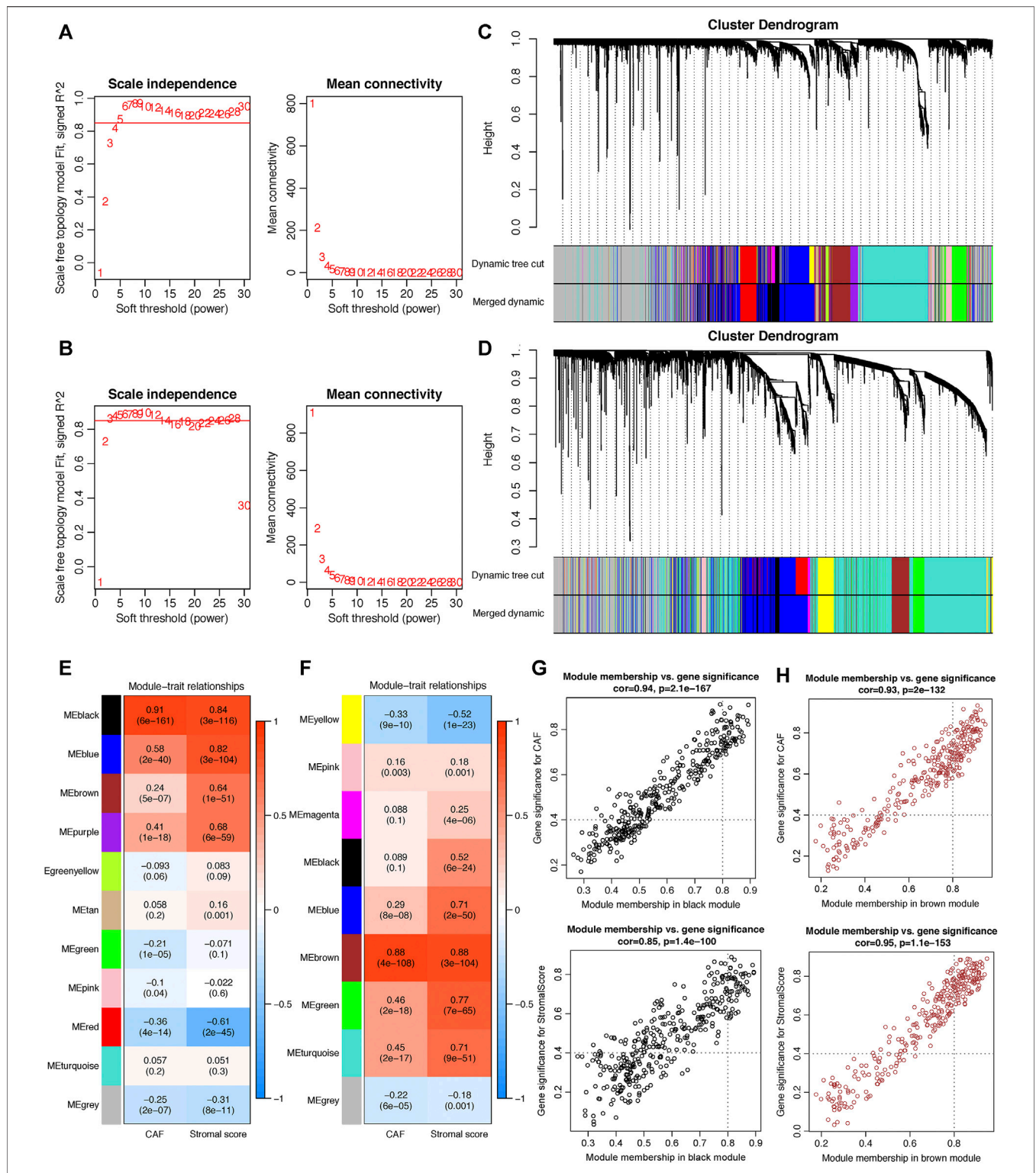


FIGURE 3 | Co-expression network constructed by WGCNA. **(A–B)** The soft-thresholding power (β) of 8 and 6 was, respectively, selected based on the scale-free topology criterion in GSE84437 **(A)** and TCGA-STAD **(B)**. **(C–D)** Clustering dendrograms showing genes with similar expression patterns were clustered into co-expression modules in GSE84437 **(C)** and TCGA-STAD **(D)**. The gray module indicates that genes were not assigned to any module. **(E–F)** Module-trait relationships revealing the correlations between each gene module eigengene and phenotype in GSE84437 **(E)** and TCGA-STAD **(F)**. **(G–H)** Scatter plots of the module membership (MM) and gene significance (GS) of each gene in the black module of GSE84437 **(G)** and the brown module of TCGA-STAD **(H)**. The horizontal axis is the correlation between the gene and co-expression module, and the vertical axis is the correlation between the gene and phenotype.

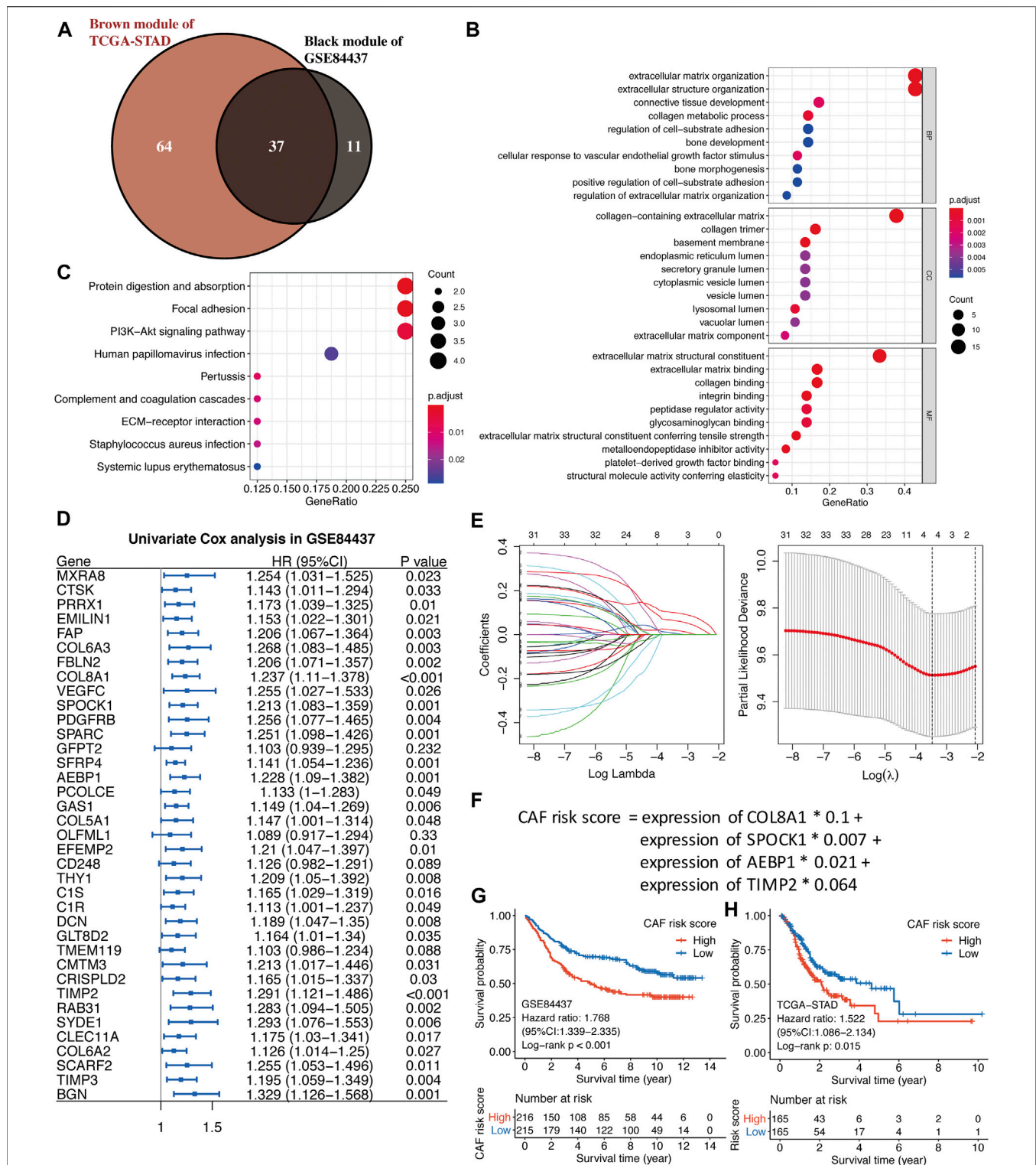


FIGURE 4 | (A) The intersection of GSE84437 black and TCGA-STAD brown module genes was presented in the Venn diagram. **(B–C)** GO analyses of the enriched biological process (BP), cellular component (CC), and molecular function (MF) terms **(B)** and KEGG pathway analysis **(C)** of the 37 genes. **(D)** Univariate Cox analysis for the screening of overall survival-associated genes in GSE84437. **(E)** Coefficient profiles of least absolute shrinkage and selection operator (LASSO) Cox regression analysis, and the adjustment parameter (lambda) was calculated based on the partial likelihood deviance with ten-fold cross validation. **(F)** Formulation of the CAF risk model. **(G,H)** Kaplan-Meier analyses identified gastric cancer patients in the high-CAF-risk group which exhibited worse overall survival in both GSE84437 **(G)** and TCGA-STAD **(H)** cohorts.

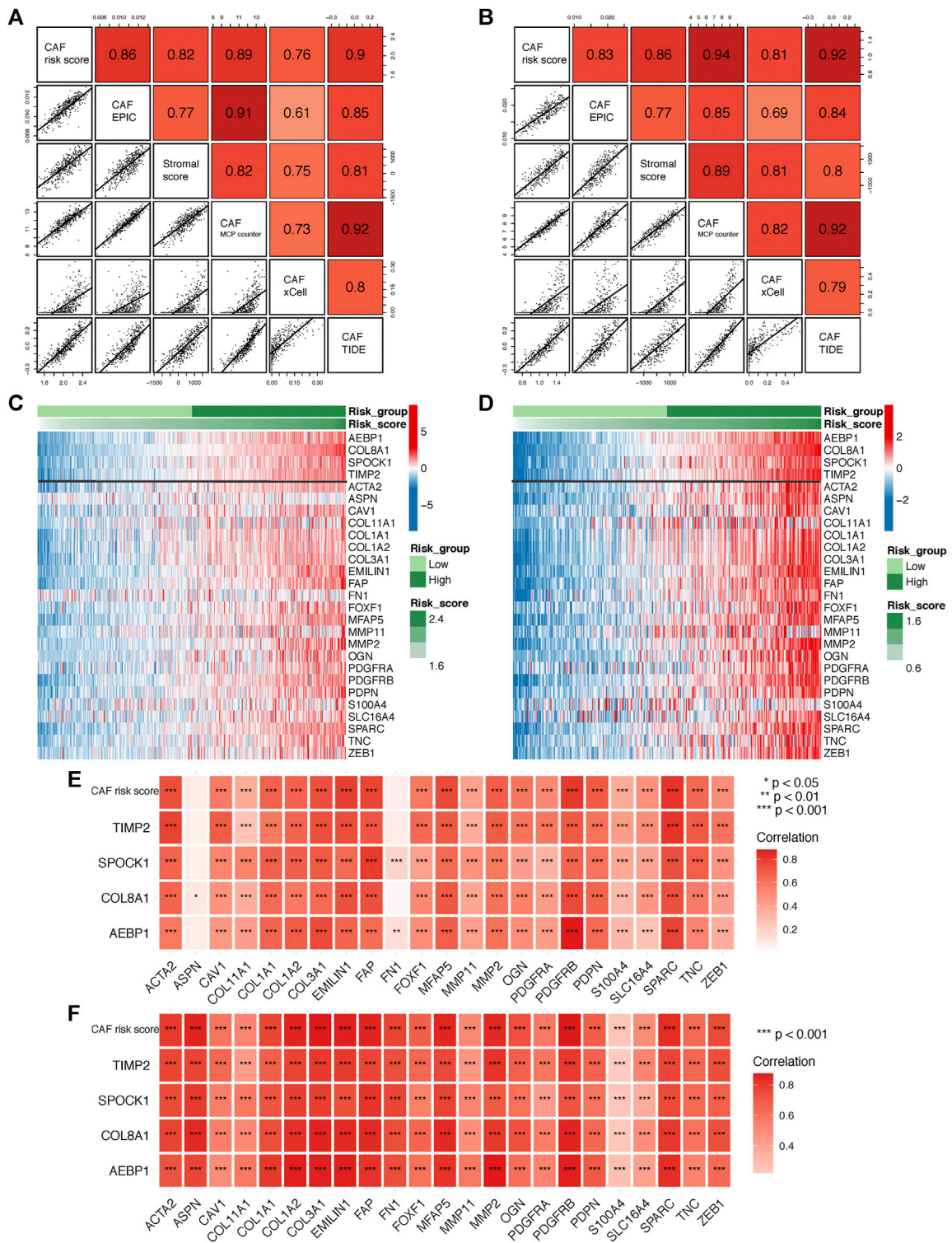
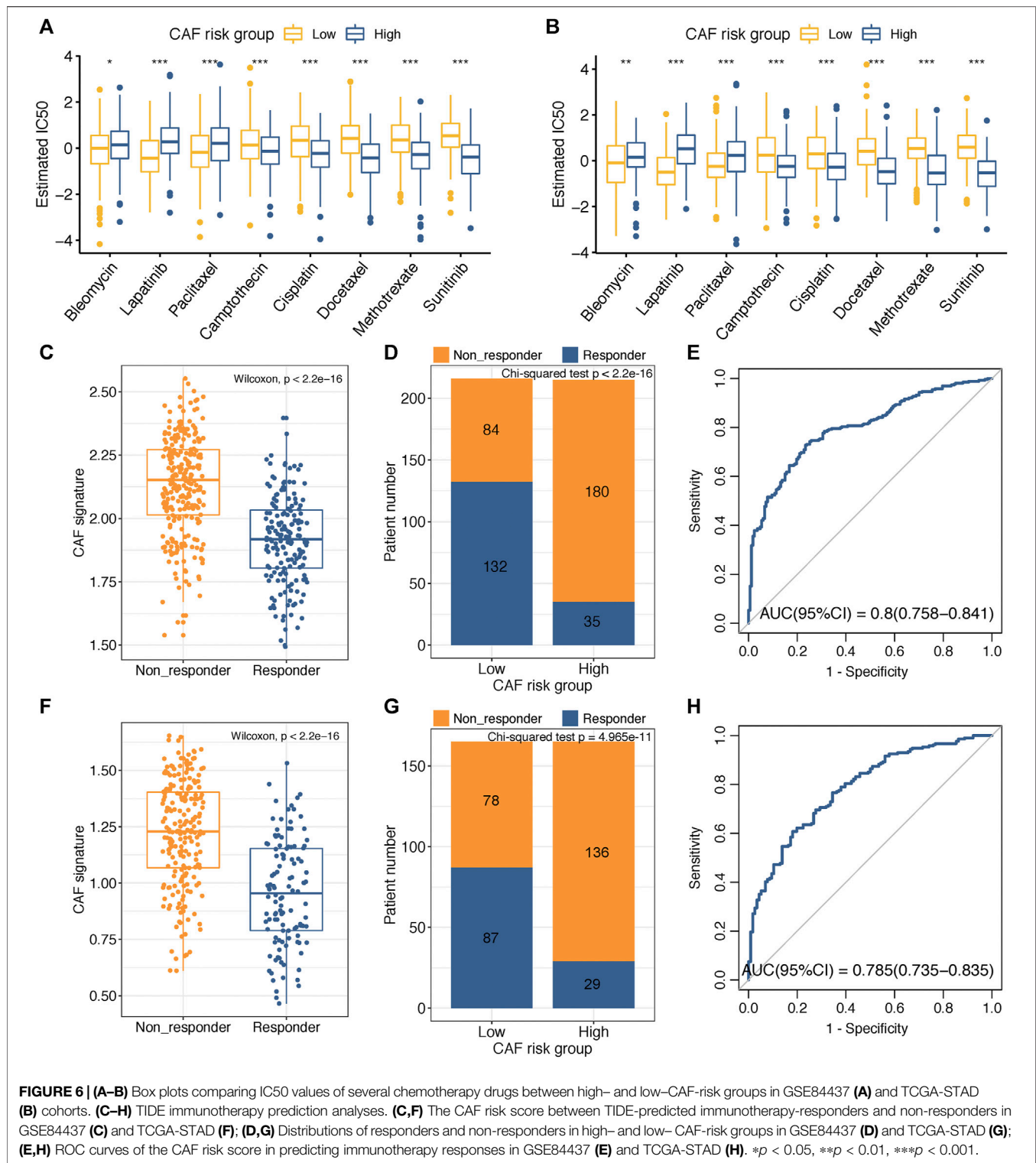


FIGURE 5 | (A–B) Spearman’s correlation analyses revealing the CAF risk score was strongly and positively correlated with stromal scores and multi-estimated CAF infiltrations in GSE84437 **(A)** and TCGA-STAD **(B)** cohorts. **(C–D)** The heat map revealing the expression patterns of CAF markers identified four CAF genes with the CAF risk score in GSE84437 **(C)** and TCGA-STAD **(D)** cohorts. **(E–F)** The CAF risk score and four signature genes were positively correlated with literature that reported CAF markers in GSE84437 **(E)** and TCGA-STAD **(F)** cohorts.



GDSC database. Wilcoxon analyses identified significant differences in IC50 values between GC patients in high- and low-CAF-risk groups, with high-CAF-risk GC patients revealing increased sensitivity to camptothecin, cisplatin, docetaxel, methotrexate, and sunitinib, while the low-CAF score subgroup was estimated to be more sensitive to bleomycin,

lapatinib, and paclitaxel in both GSE84437 **(Figure 6A)** and TCGA-STAD **(Figure 6B)** cohorts.

Immunotherapy using immune checkpoint inhibitors has brought hope to GC patients. We applied the TIDE method to assess whether the CAF risk score could serve as an immunotherapy predictor for GC patients. For GSE84437, the

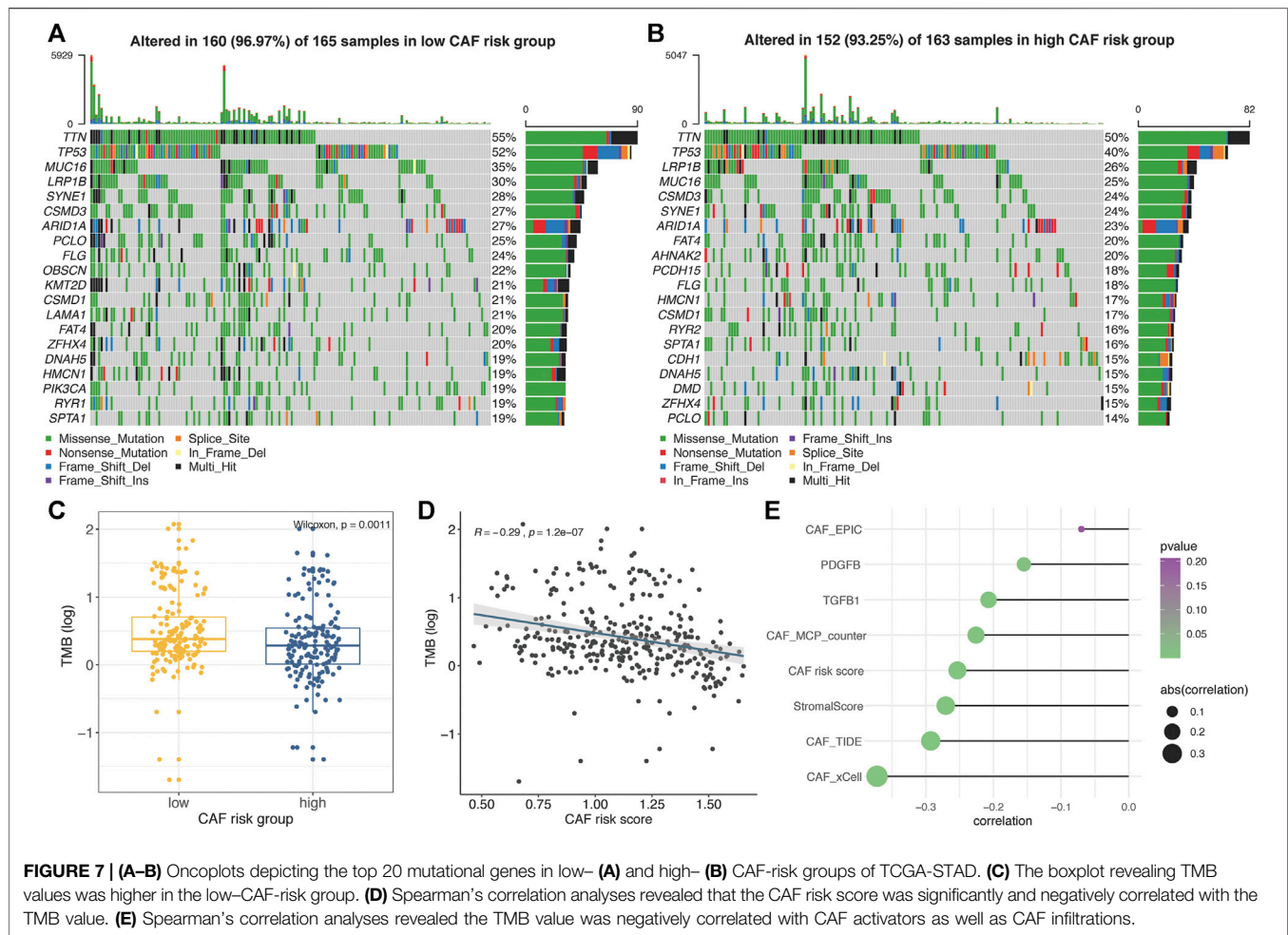


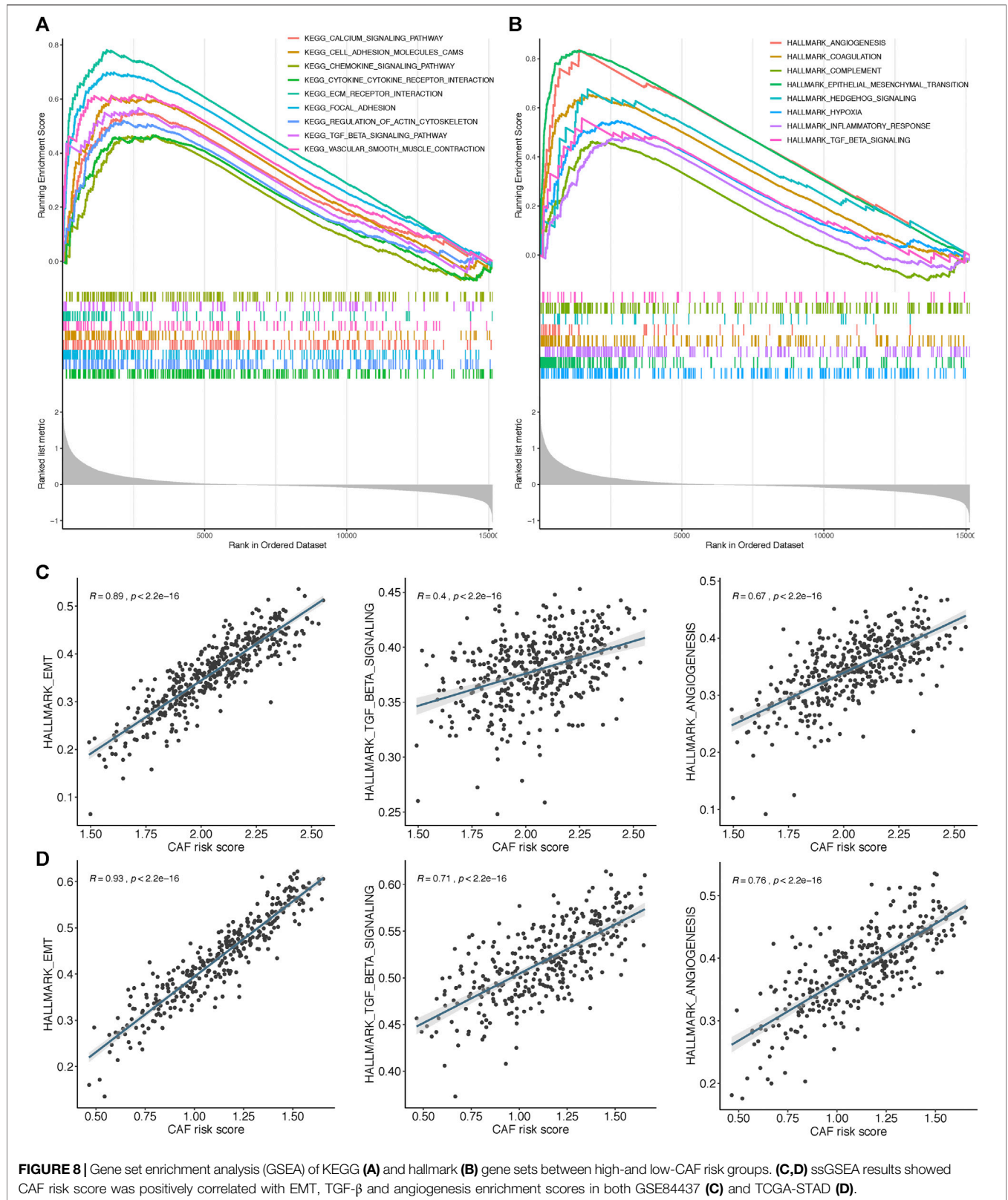
FIGURE 7 | (A–B) Oncoplots depicting the top 20 mutational genes in low– (A) and high– (B) CAF-risk groups of TCGA-STAD. **(C)** The boxplot revealing TMB values was higher in the low-CAF-risk group. **(D)** Spearman’s correlation analyses revealed that the CAF risk score was significantly and negatively correlated with the TMB value. **(E)** Spearman’s correlation analyses revealed the TMB value was negatively correlated with CAF activators as well as CAF infiltrations.

CAF score in the non-responder subgroup ($n = 264$) was significantly higher than that in the responder cohort ($n = 167$) (Wilcoxon test, $p < 2.2e-16$; **Figure 6C**); higher sensitivity to immunotherapy was observed for GC patients in the low-CAF-risk group (132/216) than that in the high-CAF-risk group (35/215) (Chi-square test, $p < 2.2e-16$; **Figure 6D**). For TCGA-STAD, the CAF score was also significantly higher in the non-responder subgroup ($n = 214$) than that in the responder cohort ($n = 116$) (Wilcoxon test, $p < 2.2e-16$; **Figure 6F**); GC patients in the low-CAF-risk group were much more sensitive to immunotherapy (87/165) than those in high-CAF risk group (29/165) (chi-square test, $p < 2.2e-16$; **Figure 6G**). Furthermore, the AUC values of 0.8 (95%CI: 0.758–0.841) in GSE84437 (**Figure 6E**) and 0.785 (95%CI: 0.735–0.835) in TCGA-STAD (**Figure 6H**) indicated the excellent performance of our CAF model for immunotherapy response predictions.

Correlation Between CAF-Related Signature and Somatic Variation

The top 20 genes with highest mutational frequencies in the low– (**Figure 7A**) and high– (**Figure 7B**) CAF-risk subgroups

were, respectively, recognized and displayed as waterfall plots. Intriguingly, several frequent mutational genes were shared in both low– and high-CAF-risk groups, including TTN, TP53, MUC16, LRP1B, SYNE1, CSMD3, ARID1A, PCLO, FLG, CSMD1, FAT4, ZFH4, DNAH5, HMCN1, and SPTA1. In addition, mutations of LAMA1, RYR1, OBSCN, KMT2D, and PIK3CA genes were more common in the low-CAF-risk group, while mutations of DMD, AHNAK2, PCDH15, RYR2, and CDH1 belonged specifically to the top 20 frequent mutational genes in the high-CAF-risk group. Subsequently, we observed that the TMB values were significantly higher in the low-CAF-score subgroup (Wilcoxon test, $p = 0.0011$, **Figure 7C**), and Spearman’s correlation analysis revealed that the CAF-risk score was significantly and negatively correlated with the TMB value (Cor = -0.29 , $p = 1.2e-07$, **Figure 7D**). Furthermore, Spearman correlation analyses also confirmed that the TMB values were negatively correlated with stromal CAF infiltrations as well as CAF-activating factors like TGF- β (Quante et al., 2011) and PDGF (Pietras et al., 2008) (**Figure 7E**), suggesting that higher TMB might be also able to intense tumor-killing effects *via* modulating a stromal fibroblast-weak local microenvironment.



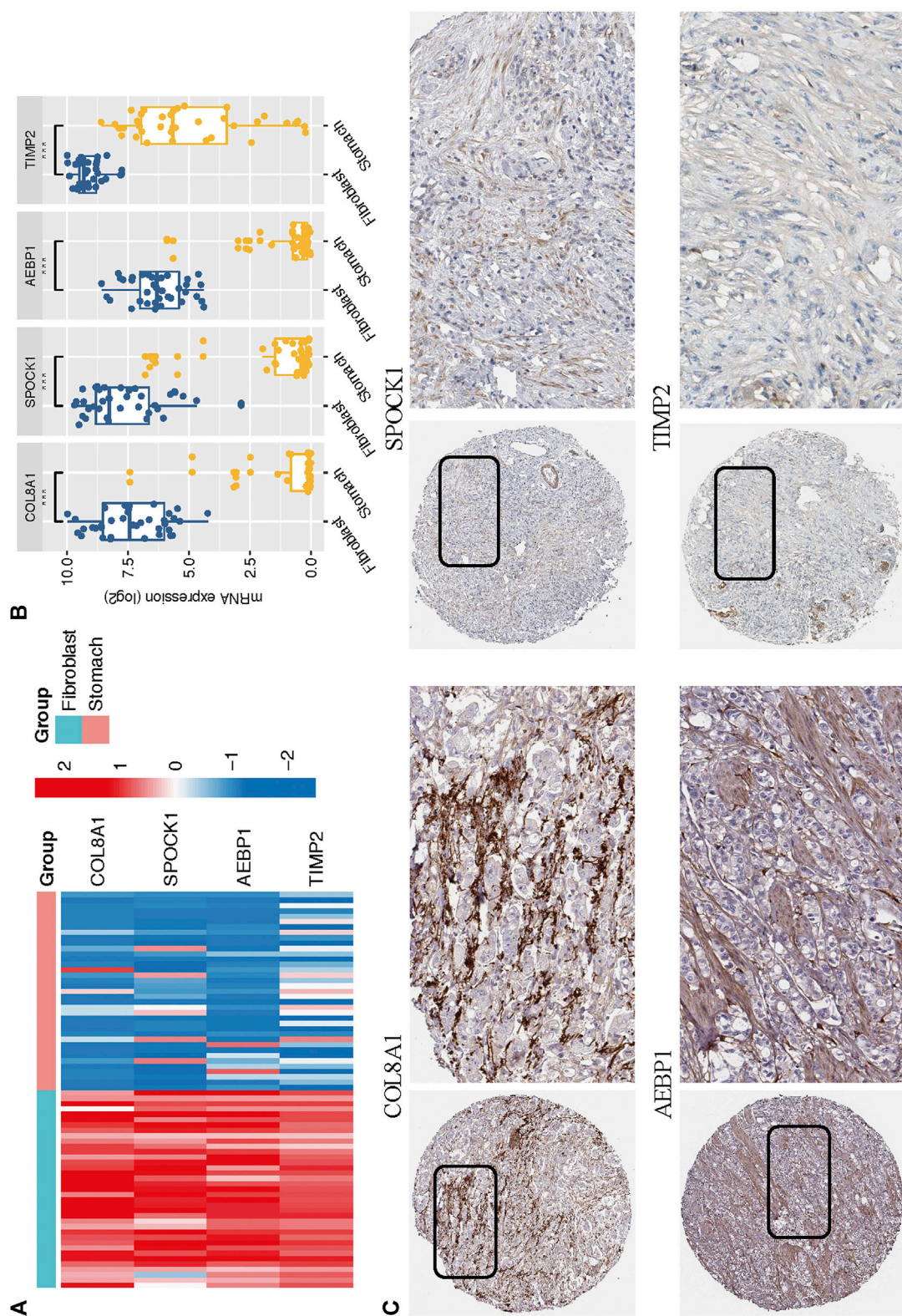


FIGURE 9 | (A–B) The mRNA expression levels of the four CAF genes in the fibroblasts and gastric cancer cell lines were illustrated in the heat map **(A)** and compared by Wilcoxon analysis **(B)**. **(C)** Protein expressions of COL8A1, SPOCK1, AEBP1, and TIMP2 in gastric cancer specimens from the Human Protein Atlas database.

GSEA of the Four-Gene CAF Signature

To further elucidate the functional enrichment of the CAF signature, GSEA was performed on the GSE84437 dataset between high- and low-CAF-risk groups. As displayed in **Figure 8A**, the major enriched KEGG signaling pathways were calcium signaling pathway, ECM receptor interaction, chemokine signaling, and transforming growth factor beta (TGF- β) signaling pathways. Additionally, genes in the high-CAF-risk group were mainly enriched in angiogenesis, epithelial-mesenchymal transition (EMT), inflammatory response, and TGF- β signaling hallmark gene sets (**Figure 8B**). Extensively, ssGSEA results also showed the CAF risk score was positively correlated with EMT, TGF- β , and angiogenesis enrichment scores in both GSE84437 (**Figure 8C**) and TCGA-STAD (**Figure 8D**).

Multidimensional Validation of Key Genes in CCLE and HPA Databases

Based on the CCLE database, we verified that the mRNA expressions of the four hub genes (COL8A1, SPOCK1, AEBP1, and TIMP2) were higher in fibroblast cell lines than those in GC cell lines (Wilcoxon test, all $p < 0.001$; **Figures 9A,B**). In addition, to determine the protein expression characteristics of these CAF signature genes, we analyzed the IHC images from the HPA database. The data demonstrated that these proteins were deeply stained in GC stroma (**Figure 9C**). These verifications implied that these four genes might be CAF-specific markers.

DISCUSSION

Gastric cancers, especially poorly and undifferentiated gastric cancers, often exhibit massive fibrosis with abundant infiltration of CAFs, which shield TME from antitumor lymphocyte infiltrations and contribute to GC progression, treatment resistance, and immunosuppression (Abe et al., 2017; Ham et al., 2019). Consistently, we observed that higher CAF and stromal scores were associated with worse OS after initial treatment in GC. Therefore, investigation of novel molecular targets in GC is pivotal for the development of stromal CAF-targeting therapies. This is the first research based on WGCNA and multiple computational algorithms to mine the mutual CAF and stromal co-expressed networks in 2 GC cohorts: GSE84437 and TCGA-STAD. By applying univariate Cox and LASSO regression algorithms, a four-gene (COL8A1, SPOCK1, AEBP1, and TIMP2) prognostic CAF model was constructed and validated. By taking the median CAF risk score as the cutoff value, we observed high-CAF-risk GC patients were more sensitive to camptothecin, cisplatin, docetaxel, methotrexate, and sunitinib. In addition, based on the TIDE online algorithm, we observed the lower CAF risk score was highly correlated with improved immunotherapeutic effects in GC patients, and higher TMB levels were observed in low-CAF-risk group STAD patients, which indicated that the CAF model could potentially serve as an immunotherapeutic stratification biomarker for GC. However, interactions between TMB and CAF

infiltrations have not been well studied to date. Our study first revealed that the TMB levels were also negatively associated with CAF activators as well as infiltrations in GC patients. It is widely acknowledged that cancer cells with a high level of mutations are easier to be recognized by the immune system, which can then strengthen the immune response and lead to improved immunotherapeutic efficacy (Miao et al., 2018). Based on our analyses, we propose another mechanism that higher TMB might also be able to intense tumor-killing effects *via* modulating a stromal fibroblast-weak local microenvironment. However, more experiments are needed to clarify the crosstalk between CAFs and TMB.

GSEA revealed that EMT, TGF- β signaling, hypoxia, and angiogenesis gene sets were highly and significantly enriched in the high-CAF-risk group; ssGSEA results also showed that the CAF risk score was positively correlated with EMT, TGF- β , and angiogenesis-enrichment scores in both two cohorts. Polarized epithelial cells gain invasive capacities through the EMT process (Huang et al., 2015), and TGF- β signaling has been reported to be responsible for the CAF activation (Yeung et al., 2013; Zheng et al., 2016; Ishimoto et al., 2017). Interactively, CAFs are capable of synergistically initiating and enhancing EMT (Bhowmick et al., 2004; Lee et al., 2006; Thiery and Sleeman, 2006; Ham et al., 2019); CAFs could also regulate and maintain the stemness of gastric cancer cells *via* TGF β signaling (Hasegawa et al., 2014). Pathological angiogenesis has been widely described as a crucial process enabling the expansion of cancerous tissues, as well as the invasion and metastasis of GC cells (Chen et al., 2004; Hoff and Machado, 2012; Forma et al., 2021). CAFs contributed dominantly to the uncontrolled angiogenesis by inducing a hypoxia TME (Kugeratski et al., 2019) and producing pro-angiogenic factors like galectin-1 (Tang et al., 2016), vascular endothelial growth factor (VEGF) (De Francesco et al., 2013), and hepatocyte growth factor (HGF) (Ding et al., 2018).

To guarantee the model's robustness and avoid over-fitting, we adopted four bioinformatics methods to quantify CAF infiltrations in GC: the EPIC method for model construction and xCell, MCP-counter, and TIDE methods for correlation verifications, and we found that our model was strongly correlated with CAF infiltrations as well as CAF markers. Meanwhile, according to the CCLE database, we further confirmed that the expressions of four identified genes were significantly higher in fibroblast cell lines, and IHC images from the HPA database also revealed higher staining of these proteins in stromal parts of GC. These results implied these genes as CAF-specific markers for GC, and our model was capable of accurately assessing CAF infiltration levels.

With respect to the four identified markers in the model, elevated expression of COL8A1 has been found in CAFs and was significantly associated with a high risk of death in head and neck squamous cell carcinoma (Lai et al., 2019). Zhang et al. identified COL8A1 as the prognostic hub gene highly correlated with Wnt2, which is elevated selectively in CAFs, and high co-expression of COL8A1 and Wnt2 was an independent adverse prognostic factor for colon cancer patients (Katoh, 2001; Zhang et al., 2020). At an epithelial cellular level, Zhou et al. reported that the knockdown of COL8A1 significantly suppressed the

proliferation and promoted the apoptosis of GC cells (Zhou et al., 2020). As for SPOCK1, studies have proved SPOCK1 as an EMT-related marker that was closely correlated with tumorigenesis and invasiveness in gastric cancer (Yan et al., 2017; Chen et al., 2018), prostate cancer (Wang et al., 2016), pancreatic cancer (Li et al., 2020), gallbladder cancer (Shu et al., 2015), and lung cancer (Miao et al., 2013). We observed that high-CAF-risk group GC patients were less sensitive to several drugs like lapatinib, and this result fits well with the finding that SPOCK1-regulated EMT derived the acquiring of lapatinib resistance in HER2-positive GC (Kim et al., 2014). Veenstra et al. identified that SPOCK1 expressed restrictively in tumor stroma, and the stromal SPOCK1 would promote pancreatic ductal adenocarcinoma invasion by mediating extracellular matrix remodeling (Veenstra et al., 2017). Sasaki et al. identified AEBP1 as a novel CAF- and EMT-related protein responsible for tumor invasiveness and metastasis in basal cell carcinoma, squamous cell carcinoma, and malignant melanoma (Sasaki et al., 2018). AEBP1 has also been reported as a pivotal proinflammatory mediator (Majdalawieh et al., 2006; Majdalawieh et al., 2007; Majdalawieh and Ro, 2010), and the overexpression of stromal AEBP1 would induce mammary tumorigenesis *via* paracrine proinflammatory signaling (Holloway et al., 2012). In addition, AEBP1 is upregulated in vascular endothelial cells and promotes tumor angiogenesis in colorectal cancer by inducing angiogenesis-related genes like AQP1 (Yorozu et al., 2020). AEBP1 has also been demonstrated as an adverse prognostic marker in GC that facilitates invasion and migration, metastasis, and EMT of GC cells *via* activating NF- κ B signaling (Liu et al., 2018). TIMP2 was previously found as a matrix metalloproteinase (MMP) inhibitor (Basu et al., 2012) that restrained cell proliferation and metastasis in GC (Johansson et al., 2010) and breast cancer (Mendes et al., 2007). However, literature on its functions in cancer is inconsistent. TIMP2 was correlated with higher pN and pM stages as well as unfavorable prognosis in GC (Alakus et al., 2010; Wang et al., 2018). Besides, TIMP2 expressed by CAFs was independently related to lower relapse-free and overall survival in breast cancer (Eiró et al., 2015; Cid et al., 2018). Nonetheless, not much of their functions are known in CAFs of GC, which necessitates further experiments of the four CAF marker's mechanisms that underline the invasiveness and metastasis, drug resistance, and immunosuppression of GC.

Some limitations should be noted in our research. First, this was a retrospective bioinformatic analysis based on two public

gene expression data; the real prognostic and therapeutic values of the CAF model should be cross-validated in multicenter and perspective data. In addition, the specific biological roles of the CAF signature biomarkers in GC should be verified through molecular and animal experiments. Nevertheless, our instructive results could serve as a framework for the future mapping of CAFs in GC.

CONCLUSION

In summary, we confirmed higher infiltration of stromal CAFs in TME was correlated with worse prognosis in GC and identified COL8A1, SPOCK1, AEBP1, and TIMP2 as novel prognostic CAF biomarkers by constructing an integrated co-expression network of infiltrated CAF and stromal scores. A CAF model comprising of these four markers was constructed and validated, which could precisely predict prognosis, CAF infiltrations, chemotherapy, and immunotherapy responses in GC patients, which might provide novel insights for targeting CAF therapeutic strategies in GC.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. These data can be found here: GDC TCGA Hub of the UCSC XENA database (<https://gdc.xenahubs.net/>); GEO datasets (<https://www.ncbi.nlm.nih.gov/geo/>): GSE84437.

AUTHOR CONTRIBUTIONS

HZ and HSL contributed to the conception and design. HZ and HYL extracted the data from the databases. HZ, HSL, and XW contributed to the data analysis and interpretation. HZ and WDD drafted the manuscript. HSL and XW revised the manuscript. XW supervised the entire study. All authors contributed to the article and approved the final manuscript.

ACKNOWLEDGMENTS

We would like to thank the TCGA and GEO databases for the availability of the data.

REFERENCES

- Abe, A., Nagatsuma, A. K., Higuchi, Y., Nakamura, Y., Yanagihara, K., and Ochiai, A. (2017). Site-Specific Fibroblasts Regulate Site-Specific Inflammatory Niche Formation in Gastric Cancer. *Gastric Cancer* 20 (1), 92–103. doi:10.1007/s10120-015-0584-y
- Alakus, H., Afriani, N., Warnecke-Eberz, U., Bollschweiler, E., Fetzner, U., Drebber, U., et al. (2010). Clinical Impact of MMP and TIMP Gene Polymorphisms in Gastric Cancer. *World J. Surg.* 34 (12), 2853–2859. doi:10.1007/s00268-010-0761-4
- Aran, D., Hu, Z., and Butte, A. J. (2017). xCell: Digitally Portraying the Tissue Cellular Heterogeneity Landscape. *Genome Biol.* 18 (1), 220. doi:10.1186/s13059-017-1349-1
- Basu, B., Correa de Sampaio, P., Mohammed, H., Fogarasi, M., Corrie, P., Watkins, N. A., et al. (2012). Inhibition of MT1-MMP Activity Using Functional Antibody Fragments Selected Against its Hemopexin Domain. *Int. J. Biochem. Cel Biol.* 44 (2), 393–403. doi:10.1016/j.biocel.2011.11.015
- Becht, E., Giraldo, N. A., Lacroix, L., Buttard, B., Elarouci, N., Petitprez, F., et al. (2016). Estimating the Population Abundance of Tissue-Infiltrating Immune and Stromal Cell Populations Using Gene Expression. *Genome Biol.* 17 (1), 218. doi:10.1186/s13059-016-1070-5

- Bhowmick, N. A., Chytil, A., Plieth, D., Gorska, A. E., Dumont, N., Shappell, S., et al. (2004). TGF- Signaling in Fibroblasts Modulates the Oncogenic Potential of Adjacent Epithelia. *Science*. 303 (5659), 848–851. doi:10.1126/science.1090922
- Chen, C.-N., Hsieh, F.-J., Cheng, Y.-M., Cheng, W.-F., Su, Y.-N., Chang, K.-J., et al. (2004). The Significance of Placenta Growth Factor in Angiogenesis and Clinical Outcome of Human Gastric Cancer. *Cancer Lett.* 213 (1), 73–82. doi:10.1016/j.canlet.2004.05.020
- Chen, D., Zhou, H., Liu, G., Zhao, Y., Cao, G., and Liu, Q. (2018). SPOCK1 Promotes the Invasion and Metastasis of Gastric Cancer Through Slug-Induced Epithelial-Mesenchymal Transition. *J. Cel. Mol. Med.* 22 (2), 797–807. doi:10.1111/jcmm.13357
- Cibulskis, K., Lawrence, M. S., Carter, S. L., Sivachenko, A., Jaffe, D., Sougnez, C., et al. (2013). Sensitive Detection of Somatic Point Mutations in Impure and Heterogeneous Cancer Samples. *Nat. Biotechnol.* 31 (3), 213–219. doi:10.1038/nbt.2514
- Cid, S., Eiro, N., Fernández, B., Sánchez, R., Andicoechea, A., Fernández-Muñiz, P. I., et al. (2018). Prognostic Influence of Tumor Stroma on Breast Cancer Subtypes. *Clin. Breast Cancer*. 18 (1), e123–e133. doi:10.1016/j.clbc.2017.08.008
- Colaprico, A., Silva, T. C., Olsen, C., Garofano, L., Cava, C., Garolini, D., et al. (2016). TCGAAbiolinks: an R/Bioconductor Package for Integrative Analysis of TCGA Data. *Nucleic Acids Res.* 44 (8), e71. doi:10.1093/nar/gkv1507
- De Francesco, E. M., Lappano, R., Santolla, M. F., Marsico, S., Caruso, A., and Maggiolini, M. (2013). HIF-1 α /GPER Signaling Mediates the Expression of VEGF Induced by Hypoxia in Breast Cancer Associated Fibroblasts (CAFs). *Breast Cancer Res.* 15 (4), R64. doi:10.1186/bcr3458
- Ding, X., Xi, W., Ji, J., Cai, Q., Jiang, J., Shi, M., et al. (2018). HGF Derived from Cancer associated Fibroblasts Promotes Vascularization in Gastric Cancer via PI3K/AKT and ERK1/2 Signaling. *Oncol. Rep.* 40 (2), 1185–1195. doi:10.3892/or.2018.6500
- Duperret, E. K., Trautz, A., Ammons, D., Perales-Puchalt, A., Wise, M. C., Yan, J., et al. (2018). Alteration of the Tumor Stroma Using a Consensus DNA Vaccine Targeting Fibroblast Activation Protein (FAP) Synergizes With Antitumor Vaccine Therapy in Mice. *Clin. Cancer Res.* 24 (5), 1190–1201. doi:10.1158/1078-0432.CCR-17-2033
- Eiró, N., Fernandez-García, B., Vázquez, J., Del Casar, J. M., González, L. O., and Vizoso, F. J. (2015). A Phenotype From Tumor Stroma Based on the Expression of Metalloproteases and Their Inhibitors, Associated with Prognosis in Breast Cancer. *Oncoimmunology*. 4 (7), e992222. doi:10.4161/2162402X.2014.992222
- Feig, C., Jones, J. O., Kraman, M., Wells, R. J. B., Deonarine, A., Chan, D. S., et al. (2013). Targeting CXCL12 From FAP-Expressing Carcinoma-Associated Fibroblasts Synergizes With Anti-PD-L1 Immunotherapy in Pancreatic Cancer. *Proc. Natl. Acad. Sci.* 110 (50), 20212–20217. doi:10.1073/pnas.1320318110
- Fiori, M. E., Di Franco, S., Villanova, L., Bianca, P., Stassi, G., and De Maria, R. (2019). Cancer-Associated Fibroblasts as Abettors of Tumor Progression at the Crossroads of EMT and Therapy Resistance. *Mol. Cancer*. 18 (1), 70. doi:10.1186/s12943-019-0994-2
- Forma, A., Tyczyńska, M., Kędzierawski, P., Gietka, K., and Sitarz, M. (2021). Gastric Carcinogenesis: a Comprehensive Review of the Angiogenic Pathways. *Clin. J. Gastroenterol.* 14 (1), 14–25. doi:10.1007/s12328-020-01295-1
- Gamradt, P., De La Fouchardière, C., and Hennino, A. (2021). Stromal Protein-Mediated Immune Regulation in Digestive Cancers. *Cancers*. 13 (1), 146. doi:10.3390/cancers13010146
- Gascard, P., and Tlsty, T. D. (2016). Carcinoma-Associated Fibroblasts: Orchestrating the Composition of Malignancy. *Genes Dev.* 30 (9), 1002–1019. doi:10.1101/gad.279737.116
- Geeleher, P., Cox, N., and Huang, R. S. (2014a). pRRophetic: an R Package for Prediction of Clinical Chemotherapeutic Response From Tumor Gene Expression Levels. *PLoS One*. 9, e107468. doi:10.1371/journal.pone.0107468
- Geeleher, P., Cox, N. J., and Huang, R. (2014b). Clinical Drug Response Can Be Predicted Using Baseline Gene Expression Levels and *In Vitro* Drug Sensitivity in Cell Lines. *Genome Biol.* 15 (3), R47. doi:10.1186/gb-2014-15-3-r47
- Ghandi, M., Huang, F. W., Jané-Valbuena, J., Kryukov, G. V., Lo, C. C., McDonald, E. R., et al. (2019). Next-Generation Characterization of the Cancer Cell Line Encyclopedia. *Nature*. 569 (7757), 503–508. doi:10.1038/s41586-019-1186-3
- Goldman, M. J., Craft, B., Hastie, M., Repčeka, K., McDade, F., Kamath, A., et al. (2020). Visualizing and Interpreting Cancer Genomics Data Via the Xena Platform. *Nat. Biotechnol.* 38 (6), 675–678. doi:10.1038/s41587-020-0546-8
- Ham, I.-H., Lee, D., and Hur, H. (2019). Role of Cancer-Associated Fibroblast in Gastric Cancer Progression and Resistance to Treatments. *J. Oncol.* 2019, 1–11. doi:10.1155/2019/6270784
- Han, C., Liu, T., and Yin, R. (2020). Biomarkers for Cancer-Associated Fibroblasts. *Biomark Res.* 8 (1), 64. doi:10.1186/s40364-020-00245-w
- Hanahan, D., and Coussens, L. M. (2012). Accessories to the Crime: Functions of Cells Recruited to the Tumor Microenvironment. *Cancer Cell*. 21 (3), 309–322. doi:10.1016/j.ccr.2012.02.022
- Hänzelmann, S., Castelo, R., and Guinney, J. (2013). GSVA: Gene Set Variation Analysis for Microarray and RNA-Seq Data. *BMC Bioinformatics*. 14 (1), 7. doi:10.1186/1471-2105-14-7
- Hasegawa, T., Yashiro, M., Nishii, T., Matsuoka, J., Fuyuhiko, Y., Morisaki, T., et al. (2014). Cancer-Associated Fibroblasts Might Sustain the Stemness of Scirrhous Gastric Cancer Cells via Transforming Growth Factor- β Signaling. *Int. J. Cancer*. 134 (8), 1785–1795. doi:10.1002/ijc.28520
- Hoff, P. M., and Machado, K. K. (2012). Role of Angiogenesis in the Pathogenesis of Cancer. *Cancer Treat. Rev.* 38 (7), 825–833. doi:10.1016/j.ctrv.2012.04.006
- Hofheinz, R.-D., al-Batran, S.-E., Hartmann, F., Hartung, G., Jäger, D., Renner, C., et al. (2003). Stromal Antigen Targeting by a Humanised Monoclonal Antibody: an Early Phase II Trial of Sibrotuzumab in Patients With Metastatic Colorectal Cancer. *Oncol. Res. Treat.* 26 (1), 44–48. doi:10.1159/000069863
- Holloway, R. W., Bogachev, O., Bharadwaj, A. G., McCluskey, G. D., Majdalawieh, A. F., Zhang, L., et al. (2012). Stromal Adipocyte Enhancer-Binding Protein (AEBP1) Promotes Mammary Epithelial Cell Hyperplasia via Proinflammatory and Hedgehog Signaling. *J. Biol. Chem.* 287 (46), 39171–39181. doi:10.1074/jbc.M112.404293
- Huang, L., Wu, R. L., and Xu, A. M. (2015). Epithelial-Mesenchymal Transition in Gastric Cancer. *Am. J. Transl Res.* 7 (11), 2141–2158.
- Ishimoto, T., Miyake, K., Nandi, T., Yashiro, M., Onishi, N., Huang, K. K., et al. (2017). Activation of Transforming Growth Factor Beta 1 Signaling in Gastric Cancer-Associated Fibroblasts Increases Their Motility, via Expression of RhoBoid 5 Homolog 2, and Ability to Induce Invasiveness of Gastric Cancer Cells. *Gastroenterology*. 153 (1), 191–204. doi:10.1053/j.gastro.2017.03.046
- Jiang, P., Gu, S., Pan, D., Fu, J., Sahu, A., Hu, X., et al. (2018). Signatures of T Cell Dysfunction and Exclusion Predict Cancer Immunotherapy Response. *Nat. Med.* 24 (10), 1550–1558. doi:10.1038/s41591-018-0136-1
- Johansson, E., Komuro, A., Iwata, C., Hagiwara, A., Fuse, Y., Watanabe, A., et al. (2010). Exogenous Introduction of Tissue Inhibitor of Metalloproteinase 2 Reduces Accelerated Growth of TGF- β -Disrupted Diffuse-Type Gastric Carcinoma. *Cancer Sci.* 101 (11), 2398–2403. doi:10.1111/j.1349-7006.2010.01688.x
- Katoh, M. (2001). Frequent Up-Regulation of WNT2 in Primary Gastric Cancer and Colorectal Cancer. *Int. J. Oncol.* 19 (5), 1003–1007. doi:10.3892/ijo.19.5.1003
- Kaur, A., Ecker, B. L., Douglass, S. M., Kugel, C. H., Webster, M. R., Almeida, F. V., et al. (2019). Remodeling of the Collagen Matrix in Aging Skin Promotes Melanoma Metastasis and Affects Immune Cell Motility. *Cancer Discov.* 9 (1), 64–81. doi:10.1158/2159-8290.CD-18-0193
- Kim, H.-P., Han, S.-W., Song, S.-H., Jeong, E.-G., Lee, M.-Y., Hwang, D., et al. (2014). Testican-1-Mediated Epithelial-Mesenchymal Transition Signaling Confers Acquired Resistance to Lapatinib in HER2-Positive Gastric Cancer. *Oncogene*. 33 (25), 3334–3341. doi:10.1038/onc.2013.285
- Kojima, Y., Acar, A., Eaton, E. N., Mellody, K. T., Scheel, C., Ben-Porath, I., et al. (2010). Autocrine TGF- and Stromal Cell-Derived Factor-1 (SDF-1) Signaling Drives the Evolution of Tumor-Promoting Mammary Stromal Myofibroblasts. *Proc. Natl. Acad. Sci.* 107 (46), 20009–20014. doi:10.1073/pnas.1013805107
- Kraman, M., Bambrough, P. J., Arnold, J. N., Roberts, E. W., Magiera, L., Jones, J. O., et al. (2010). Suppression of Antitumor Immunity by Stromal Cells Expressing Fibroblast Activation Protein-. *Science*. 330 (6005), 827–830. doi:10.1126/science.1195300
- Kugeratski, F. G., Atkinson, S. J., Neilson, L. J., Lilla, S., Knight, J. R. P., Serneels, J., et al. (2019). Hypoxic Cancer-Associated Fibroblasts Increase NCBP2-AS2/

- HIAR to Promote Endothelial Sprouting Through Enhanced VEGF Signaling. *Sci. Signal.* 12 (567), eaan8247. doi:10.1126/scisignal.aan8247
- Lai, S. L., Tan, M. L., Hollows, R. J., Robinson, M., Ibrahim, M., Margielewska, S., et al. (2019). Collagen Induces a More Proliferative, Migratory and Chemoresistant Phenotype in Head and Neck Cancer via DDR1. *Cancers*. 11 (11), 1766. doi:10.3390/cancers11111766
- Lakins, M. A., Ghorani, E., Munir, H., Martins, C. P., and Shields, J. D. (2018). Cancer-Associated Fibroblasts Induce Antigen-Specific Deletion of CD8 + T Cells to Protect Tumour Cells. *Nat. Commun.* 9 (1), 948. doi:10.1038/s41467-018-03347-0
- Langfelder, P., and Horvath, S. (2008). WGCNA: an R Package for Weighted Correlation Network Analysis. *BMC bioinformatics*. 9, 559. doi:10.1186/1471-2105-9-559
- Lee, J.-H., Kim, H.-I., Kim, M. G., Ha, T. K., Jung, M.-S., and Kwon, S. J. (2016). Recurrence of Gastric Cancer in Patients Who Are Disease-free for More Than 5 Years After Primary Resection. *Surgery*. 159 (4), 1090–1098. doi:10.1016/j.surg.2015.11.002
- Lee, T. K., Poon, R. T. P., Yuen, A. P., Ling, M. T., Kwok, W. K., Wang, X. H., et al. (2006). Twist Overexpression Correlates With Hepatocellular Carcinoma Metastasis Through Induction of Epithelial-Mesenchymal Transition. *Clin. Cancer Res.* 12 (18), 5369–5376. doi:10.1158/1078-0432.ccr-05-2722
- Li, J., Guan, J., Long, X., Wang, Y., and Xiang, X. (2016). Mir-1-Mediated Paracrine Effect of Cancer-Associated Fibroblasts on Lung Cancer Cell Proliferation and Chemoresistance. *Oncol. Rep.* 35 (6), 3523–3531. doi:10.3892/or.2016.4714
- Li, J., Ke, J., Fang, J., and Chen, J. P. (2020). A Potential Prognostic Marker and Therapeutic Target: SPOCK1 Promotes the Proliferation, Metastasis, and Apoptosis of Pancreatic Ductal Adenocarcinoma Cells. *J. Cell Biochem.* 121 (1), 743–754. doi:10.1002/jcb.29320
- Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J. P., and Tamayo, P. (2015). The Molecular Signatures Database Hallmark Gene Set Collection. *Cel Syst.* 1 (6), 417–425. doi:10.1016/j.cels.2015.12.004
- Liu, B., Chen, X., Zhan, Y., Wu, B., and Pan, S. (2021a). Identification of a Gene Signature for Renal Cell Carcinoma-Associated Fibroblasts Mediating Cancer Progression and Affecting Prognosis. *Front. Cel Dev. Biol.* 8, 604627. doi:10.3389/fcell.2020.604627
- Liu, B., Zhan, Y., Chen, X., Hu, X., Wu, B., and Pan, S. (2021b). Weighted Gene Co-Expression Network Analysis Can Sort Cancer-Associated Fibroblast-Specific Markers Promoting Bladder Cancer Progression. *J. Cel Physiol.* 236 (2), 1321–1331. doi:10.1002/jcp.29939
- Liu, J.-Y., Jiang, L., Liu, J.-J., He, T., Cui, Y.-H., Qian, F., et al. (2018). AEBP1 Promotes Epithelial-Mesenchymal Transition of Gastric Cancer Cells by Activating the NF-Kb Pathway and Predicts Poor Outcome of the Patients. *Sci. Rep.* 8 (1), 11955. doi:10.1038/s41598-018-29878-6
- Lotti, F., Jarrar, A. M., Pai, R. K., Hitomi, M., Lathia, J., Mace, A., et al. (2013). Chemotherapy Activates Cancer-Associated Fibroblasts to Maintain Colorectal Cancer-Initiating Cells by IL-17A. *J. Exp. Med.* 210 (13), 2851–2872. doi:10.1084/jem.20131195
- Ma, H.-Y., Liu, X.-Z., and Liang, C.-M. (2016). Inflammatory Microenvironment Contributes to Epithelial-Mesenchymal Transition in Gastric Cancer. *World J. Gastroenterol.* 22 (29), 6619–6628. doi:10.3748/wjg.v22.i29.6619
- Majdalawieh, A., and Ro, H.-S. (2010). PPAR γ 1 and LXR α Face a New Regulator of Macrophage Cholesterol Homeostasis and Inflammatory Responsiveness, AEBP1. *Nucl. Recept Signal.* 8, nrs.08004–e004. doi:10.1621/nrs.08004
- Majdalawieh, A., Zhang, L., Fuki, I. V., Rader, D. J., and Ro, H.-S. (2006). Adipocyte Enhancer-Binding Protein 1 Is a Potential Novel Atherogenic Factor Involved in Macrophage Cholesterol Homeostasis and Inflammation. *Proc. Natl. Acad. Sci. U S A.* 103 (7), 2346–2351. doi:10.1073/pnas.0508139103
- Majdalawieh, A., Zhang, L., and Ro, H.-S. (2007). Adipocyte Enhancer-Binding Protein-1 Promotes Macrophage Inflammatory Responsiveness by Up-Regulating NF-Kb via I κ B α Negative Regulation. *MBoC.* 18 (3), 930–942. doi:10.1091/mbc.e06-03-0217
- Mayakonda, A., Lin, D.-C., Assenov, Y., Plass, C., and Koefler, H. P. (2018). Maftools: Efficient and Comprehensive Analysis of Somatic Variants in Cancer. *Genome Res.* 28 (11), 1747–1756. doi:10.1101/gr.239244.118
- Mendes, O., Kim, H.-T., Lungu, G., and Stoica, G. (2007). MMP2 Role in Breast Cancer Brain Metastasis Development and its Regulation by TIMP2 and ERK1/2. *Clin. Exp. Metastasis.* 24 (5), 341–351. doi:10.1007/s10585-007-9071-0
- Miao, D., Margolis, C. A., Vokes, N. I., Liu, D., Taylor-Weiner, A., Wankowicz, S. M., et al. (2018). Genomic Correlates of Response to Immune Checkpoint Blockade in Microsatellite-Stable Solid Tumors. *Nat. Genet.* 50 (9), 1271–1281. doi:10.1038/s41588-018-0200-2
- Miao, L., Wang, Y., Xia, H., Yao, C., Cai, H., and Song, Y. (2013). SPOCK1 Is a Novel Transforming Growth Factor- β Target Gene That Regulates Lung Cancer Cell Epithelial-Mesenchymal Transition. *Biochem. Biophysical Res. Commun.* 440 (4), 792–797. doi:10.1016/j.bbrc.2013.10.024
- Monteran, L., and Erez, N. (2019). The Dark Side of Fibroblasts: Cancer-Associated Fibroblasts as Mediators of Immunosuppression in the Tumor Microenvironment. *Front. Immunol.* 10, 1835. doi:10.3389/fimmu.2019.01835
- Narra, K., Mullins, S. R., Lee, H.-O., Strzemkowski-Brun, B., Magalong, K., Christiansen, V. J., et al. (2007). Phase II Trial of Single Agent Val-boroPro (Talabostat) Inhibiting Fibroblast Activation Protein in Patients With Metastatic Colorectal Cancer. *Cancer Biol. Ther.* 6 (11), 1691–1699. doi:10.4161/cbt.6.11.4874
- Pietras, K., Pahler, J., Bergers, G., and Hanahan, D. (2008). Functions of Paracrine PDGF Signaling in the Proangiogenic Tumor Stroma Revealed by Pharmacological Targeting. *Plos Med.* 5 (1), e19. doi:10.1371/journal.pmed.0050019
- Quail, D. F., and Joyce, J. A. (2013). Microenvironmental Regulation of Tumor Progression and Metastasis. *Nat. Med.* 19 (11), 1423–1437. doi:10.1038/nm.3394
- Quante, M., Tu, S. P., Tomita, H., Gonda, T., Wang, S. S. W., Takashi, S., et al. (2011). Bone Marrow-Derived Myofibroblasts Contribute to the Mesenchymal Stem Cell Niche and Promote Tumor Growth. *Cancer Cell.* 19 (2), 257–272. doi:10.1016/j.ccr.2011.01.020
- Racle, J., de Jonge, K., Baumgaertner, P., Speiser, D. E., and Gfeller, D. (2017). Simultaneous Enumeration of Cancer and Immune Cell Types From Bulk Tumor Gene Expression Data. *eLife.* 6, e26476. doi:10.7554/eLife.26476
- Sasaki, K., Sugai, T., Ishida, K., Osakabe, M., Amano, H., Kimura, H., et al. (2018). Analysis of Cancer-Associated Fibroblasts and the Epithelial-Mesenchymal Transition in Cutaneous Basal Cell Carcinoma, Squamous Cell Carcinoma, and Malignant Melanoma. *Hum. Pathol.* 79, 1–8. doi:10.1016/j.humpath.2018.03.006
- Shu, Y.-J., Weng, H., Ye, Y.-Y., Hu, Y.-P., Bao, R.-F., Cao, Y., et al. (2015). SPOCK1 as a Potential Cancer Prognostic Marker Promotes the Proliferation and Metastasis of Gallbladder Cancer Cells by Activating the PI3K/AKT Pathway. *Mol. Cancer.* 14 (1), 12. doi:10.1186/s12943-014-0276-y
- Simon, N., Friedman, J., Hastie, T., and Tibshirani, R. (2011). Regularization Paths for Cox's Proportional Hazards Model via Coordinate Descent. *J. Stat. Soft.* 39 (5), 1–13. doi:10.18637/jss.v039.i05
- Sturm, G., Finotello, F., and List, M. (2020). “Immuneconv: An R Package for Unified Access to Computational Methods for Estimating Immune Cell Fractions From Bulk RNA-Sequencing Data,” in *Bioinformatics for Cancer Immunotherapy: Methods and Protocols*. Editor S. Boegel (New York, NY: Springer US), 223–232. doi:10.1007/978-1-0716-0327-7_16
- Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., et al. (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA A. Cancer J. Clin.* 71 (3), 209–249. doi:10.3322/caac.21660
- Tang, D., Gao, J., Wang, S., Ye, N., Chong, Y., Huang, Y., et al. (2016). Cancer-Associated Fibroblasts Promote Angiogenesis in Gastric Cancer Through Galectin-1 Expression. *Tumor Biol.* 37 (2), 1889–1899. doi:10.1007/s13277-015-3942-9
- Thiery, J. P., and Sleeman, J. P. (2006). Complex Networks Orchestrate Epithelial-Mesenchymal Transitions. *Nat. Rev. Mol. Cel Biol.* 7 (2), 131–142. doi:10.1038/nrm1835
- Thrift, A. P., and El-Serag, H. B. (2020). Burden of Gastric Cancer. *Clin. Gastroenterol. Hepatol.* 18 (3), 534–542. doi:10.1016/j.cgh.2019.07.045
- Tommelein, J., Verset, L., Boterberg, T., Demetter, P., Bracke, M., and De Wever, O. (2015). Cancer-Associated Fibroblasts Connect Metastasis-Promoting Communication in Colorectal Cancer. *Front. Oncol.* 5, 63. doi:10.3389/fonc.2015.00063
- Uhlén, M., Fagerberg, L., Hallström, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., et al. (2015). Tissue-Based Map of the Human Proteome. *Science.* 347 (6220), 1260419. doi:10.1126/science.1260419

- Veenstra, V. L., Damhofer, H., Waasdorp, C., Steins, A., Kocher, H. M., Medema, J. P., et al. (2017). Stromal SPOCK1 Supports Invasive Pancreatic Cancer Growth. *Mol. Oncol.* 11 (8), 1050–1064. doi:10.1002/1878-0261.12073
- Wagner, G. P., Kin, K., and Lynch, V. J. (2012). Measurement of mRNA Abundance Using RNA-Seq Data: RPKM Measure Is Inconsistent Among Samples. *Theor. Biosci.* 131 (4), 281–285. doi:10.1007/s12064-012-0162-3
- Wang, W., Zhang, Y., Liu, M., Wang, Y., Yang, T., Li, D., et al. (2018). TIMP2 Is a Poor Prognostic Factor and Predicts Metastatic Biological Behavior in Gastric Cancer. *Sci. Rep.* 8 (1), 9629. doi:10.1038/s41598-018-27897-x
- Wang, Z., Yao, Y.-T., Xu, H., Chen, Y.-B., Gu, M., Cai, Z.-K., et al. (2016). SPOCK1 Promotes Tumor Growth and Metastasis in Human Prostate Cancer. *Drug Des. Devel Ther.* 10, 2311–2321. doi:10.2147/DDDT.S91321
- Wu, X., Tao, P., Zhou, Q., Li, J., Yu, Z., Wang, X., et al. (2017). IL-6 Secreted by Cancer-Associated Fibroblasts Promotes Epithelial-Mesenchymal Transition and Metastasis of Gastric Cancer via JAK2/STAT3 Signaling Pathway. *Oncotarget.* 8 (13), 20741–20750. doi:10.18632/oncotarget.15119
- Yan, L., Sun, K., Liu, Y., Liang, J., Cai, K., and Gui, J. (2017). MiR-129-5p Influences the Progression of Gastric Cancer Cells Through Interacting With SPOCK1. *Tumour Biol.* 39 (6), 101042831770691. doi:10.1177/1010428317706916
- Yang, W., Soares, J., Greninger, P., Edelman, E. J., Lightfoot, H., Forbes, S., et al. (2013). Genomics of Drug Sensitivity in Cancer (GDSC): a Resource for Therapeutic Biomarker Discovery in Cancer Cells. *Nucleic Acids Res.* 41 (D1), D955–D961. doi:10.1093/nar/gks1111
- Yarchoan, M., Hopkins, A., and Jaffee, E. M. (2017). Tumor Mutational Burden and Response Rate to PD-1 Inhibition. *N. Engl. J. Med.* 377 (25), 2500–2501. doi:10.1056/NEJMc1713444
- Yeung, T.-L., Leung, C. S., Wong, K.-K., Samimi, G., Thompson, M. S., Liu, J., et al. (2013). TGF- β Modulates Ovarian Cancer Invasion by Upregulating CAF-Derived Versican in the Tumor Microenvironment. *Cancer Res.* 73 (16), 5016–5028. doi:10.1158/0008-5472.CAN-13-0023
- Yoon, S.-J., Park, J., Shin, Y., Choi, Y., Park, S. W., Kang, S.-G., et al. (2020). Deconvolution of Diffuse Gastric Cancer and the Suppression of CD34 on the BALB/c Nude Mice Model. *BMC cancer.* 20 (1), 314. doi:10.1186/s12885-020-06814-4
- Yorozu, A., Yamamoto, E., Niinuma, T., Tsuyada, A., Maruyama, R., Kitajima, H., et al. (2020). Upregulation of Adipocyte Enhancer-Binding Protein 1 in Endothelial Cells Promotes Tumor Angiogenesis in Colorectal Cancer. *Cancer Sci.* 111 (5), 1631–1644. doi:10.1111/cas.14360
- Yoshihara, K., Shahmoradgoli, M., Martínez, E., Vegesna, R., Kim, H., Torres-García, W., et al. (2013). Inferring Tumour Purity and Stromal and Immune Cell Admixture from Expression Data. *Nat. Commun.* 4, 2612. doi:10.1038/ncomms3612
- Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). ClusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *OMICS: A J. Integr. Biol.* 16 (5), 284–287. doi:10.1089/omi.2011.0118
- Zhang, L., Jiang, X., Li, Y., Fan, Q., Li, H., Jin, L., et al. (2020). Clinical Correlation of Wnt2 and COL8A1 With Colon Adenocarcinoma Prognosis. *Front. Oncol.* 10, 1504. doi:10.3389/fonc.2020.01504
- Zheng, L., Xu, C., Guan, Z., Su, X., Xu, Z., Cao, J., et al. (2016). Galectin-1 Mediates TGF- β -Induced Transformation From Normal Fibroblasts Into Carcinoma-Associated Fibroblasts and Promotes Tumor Progression in Gastric Cancer. *Am. J. Transl Res.* 8 (4), 1641–1658.
- Zhou, J., Song, Y., Gan, W., Liu, L., Chen, G., Chen, Z., et al. (2020). Upregulation of COL8A1 Indicates Poor Prognosis Across Human Cancer Types and Promotes the Proliferation of Gastric Cancer Cells. *Oncol. Lett.* 20 (4), 34. doi:10.3892/ol.2020.11895

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.

Copyright © 2021 Zheng, Liu, Li, Dou and Wang. 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.