



## Research article

# Gastrointestinal pan-cancer landscape of tumor matrix heterogeneity identifies biologically distinct matrix stiffness subtypes predicting prognosis and chemotherapy efficacy



Yumei Ning<sup>a,b,1</sup>, Kun Lin<sup>a,1</sup>, Jun Fang<sup>a,1</sup>, Yang Ding<sup>a,b</sup>, Zhang Zhang<sup>a,b</sup>, Xiaojia Chen<sup>a,b</sup>, Qiu Zhao<sup>a,b,\*</sup>, Haizhou Wang<sup>a,b,\*</sup>, Fan Wang<sup>a,b,\*</sup>

<sup>a</sup> Department of Gastroenterology, Zhongnan Hospital of Wuhan University, Wuhan, China

<sup>b</sup> Hubei Clinical Center and Key Lab of Intestinal and Colorectal Diseases, Wuhan, China

## ARTICLE INFO

## Article history:

Received 15 January 2023

Received in revised form 18 April 2023

Accepted 18 April 2023

Available online 20 April 2023

## Keywords:

Gastrointestinal cancer

Extracellular matrix

Matrix stiffness

Cancer subtype

Chemotherapy

## ABSTRACT

Gastrointestinal (GI) cancers are a heterogeneous group of primary solid tumors, arising in GI tract from the esophagus to rectum. Matrix stiffness (MS) is a critical physical factor for cancer progression; however, its importance in tumor progression remains to be comprehensively recognized. Herein, we conducted a comprehensive pan-cancer analysis of MS subtypes across seven GI-cancer types. Using unsupervised clustering based on literature-derived MS-specific pathway signatures, the GI-tumor samples were divided into three MS subtypes, termed as the Soft, Mixed and Stiff. Then, distinct prognoses, biological features, tumor microenvironments and mutation landscapes among three MS subtypes were revealed. The Stiff tumor subtype was associated with the poorest prognosis, the most malignant biological behaviors, and the immunosuppressive tumor stromal microenvironment. Furthermore, multiple machine learning algorithms were used to develop an 11-gene MS-signature to identify the MS subtypes of GI-cancer and predict chemotherapy sensitivity, which were further validated in two external GI-cancer cohorts. This novel MS-based classification on GI-cancers could enhance our understanding of the important role of MS in tumor progression, and may have implications for the optimization of individualized cancer management.

© 2023 Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## 1. Introduction

The surface of human gastrointestinal tract is covered with substantial epithelium, undergoing the most rapid turnover in the body [1]. As a result, gastrointestinal (GI) cancers are among the most frequent malignancies, responsible for roughly half of all cancer-related deaths [2,3]. GI-cancer include malignancies arising in the esophagus, stomach, liver and bile ducts, gallbladder, pancreas, the small intestine, colon and rectum. Differences exist in the biological and clinical features of these cancers, which may be the result of different cell and tissue origins [4]. Nevertheless, it cannot be denied that important parallels exist between GI malignancies,

and it is necessary to understand these diseases as a whole [4]. Over the past few decades, growing knowledge has been obtained on the molecular events that contribute to the development of GI-cancers. With the gradual maturity of genomic analysis in GI-cancers, significant progress has been made in transcriptome-based subtyping [5]. Molecular subtypes are highly interconnected across various GI-cancers, and many of which can be identified in spite of the organ of origin [4]. Subtyping GI-cancer based on transcriptome highlights the molecular properties associated with tumor development and integrates many factors that contribute to tumor heterogeneity [6,7].

A tumor is not simply a group of cancer cells, but rather interacts closely with the immune cells, stromal cells and extracellular matrix (ECM), which together form the major construct of the tumor microenvironment (TME) [8,9]. Based on the infiltration of immune cells, subtyping tumor into “hot” and “cold” categories has been advocated and widely accepted due to their distinct tumor immune microenvironment and biological consequences [10]. In contrast, researches about the role of stromal cells and ECM in TME are a

\* Corresponding authors at: Department of Gastroenterology, Zhongnan Hospital of Wuhan University, Wuhan, China

E-mail addresses: [qiuzhao@whu.edu.cn](mailto:qiuzhao@whu.edu.cn) (Q. Zhao),

[2016203030002@whu.edu.cn](mailto:2016203030002@whu.edu.cn) (H. Wang), [2020283030092@whu.edu.cn](mailto:2020283030092@whu.edu.cn) (F. Wang).

<sup>1</sup> Yumei Ning, Kun Lin and Jun Fang contributed equally to this work.

relative minority in the past. Nonetheless, accumulated evidence suggests that the changes of physical factors in carcinogenesis were critical for the progression of cancers [11]. ECM stress, also called matrix stiffness (MS), is one of critical physical factors in the tumor stromal microenvironment [12]. Normal tissues had an ECM with suitable MS that is a three-dimensional, non-cellular structure, usually composed of various proteins including collagens, glycoproteins, and matrix-associated proteins, providing structural and biochemical support for surrounding cells [13,14]. However, the tumor matrix often appear alterations in the density and composition, usually trending towards stiffening and rigidity [15,16]. In many types of GI-cancers, such as colorectal cancer (CRC) and pancreatic cancer, the TME exhibits higher MS than the normal tissue [17,18]. ECM-related gene signatures are reportedly associated with poor prognosis in some GI-cancers [19–21]. Additionally, recent studies in animal models reported that stiffened tumor matrix may reduce the efficacy of chemotherapy by interfering with the distribution of chemotherapeutic agents [22]. However, it remains unclear how the MS differs from various GI-cancers and the adjacent normal tissues, and whether a “stiff” or “soft” tumor matrix can inform diagnostic and therapeutic strategies in GI-cancer. Therefore, a more extensive examination of the tumor MS and the role it plays in GI-cancer is needed.

In this study, we conducted a comprehensive GI pan-cancer analysis of MS subsets among 7 types of bulk GI-cancers using RNA-sequencing and microarray data. On the basis of literature-derived matrix-specific signatures, three MS subtypes were identified, and we advocated to term them as “Soft”, “Mixed”, and “Stiff” classes. Subsequently, we revealed the landscape of hierarchical tumor matrix across different MS subtypes of GI-cancer, and determined the potential molecular and genomic features that involved in tumor matrix stiffening. Finally, a machine learning-based gene signature was developed to predict the tumor MS and chemotherapy efficacy.

## 2. Materials and methods

### 2.1. Data collection and preprocessing

Overall, seven types of GI-cancer from The Cancer Genome Atlas (TCGA) were enrolled in the study, including CHOL (cholangiocarcinoma), COAD (colon adenocarcinoma), ESCA (esophageal carcinoma), LIHC (liver hepatocellular carcinoma), PAAD (pancreatic adenocarcinoma), READ (rectum adenocarcinoma) and STAD (stomach adenocarcinoma). The transcriptomic RNA-sequencing data were downloaded from the USCS XENA portal <https://xena.ucsc.edu/> as FPKM units and converted through the  $\log_2(x + 1)$  method. The corresponding clinical information were downloaded from TCGA database (<http://cancergenome.nih.gov/>). These seven GI-cancer cohorts from TCGA were used as the main analytical data and the training set to develop the Matrix Stiffness Score (MS-score). Additionally, several microarray datasets included six GI-cancer types were used as the secondary analytical data and the validation set, including GSE39582 (COAD), GSE53625 (esophageal squamous cell carcinoma, ESCC), GSE76427 (hepatocellular carcinoma, HCC), GSE87211 (READ), GSE66229 (gastric cancer, GC), and E-MTAB-6134 (pancreatic ductal adenocarcinoma, PDAC). The transcriptome data and clinical information of the E-MTAB-6134 was downloaded from ArrayExpress (<https://www.ebi.ac.uk/biostudies/arrayexpress>). The expression profiles and clinical data of the other five microarray cohorts were obtained from the Gene Expression Omnibus (GEO) database (<http://www.ncbi.nlm.nih.gov/geo/>). These six microarray datasets were merged into a pooled Microarray cohort, and batch effects were removed using the combat function in the “sva” R package. Principal components analysis was used to detect the results by ‘prcomp’ R function (Fig. S1). After excluding samples without complete clinical information, we eventually enrolled 1685

patients with GI-cancers from TCGA and 1640 patients from the microarray databases.

### 2.2. Identification of subtypes of GI-tumors based on MS-specific profile

We manually collected 12 knowledge-based gene expression signatures associated with tumor MS from published literature [23–27] and Molecular Signatures Database (MSigDB, <http://www.gsea-msigdb.org/gsea/index.jsp>) (Table S1). The relative activation scores of each signature were calculated by single-sample gene set enrichment analysis (ssGSEA) via the “GSVA” R packages. The ssGSEA scores of 12 tumor MS-specific signatures from all TCGA datasets were combined and analyzed together. Similarly, the scores from the microarray datasets were also combined. Next, based on the MS-specific signature scores, unsupervised clustering was performed to classify GI-cancer patients into different clusters, termed by MS subtypes. The unsupervised consensus clustering was conducted by the “ConsensusClusterPlus” R package (parameters: reps = 100, pltem = 0.8, pFeature = 1). The average of the 12 ssGSEA scores called “Matrix stiffness degree” (MS degree) was used to roughly represent the overall MS of each tumor sample for subsequent analysis.

### 2.3. Pathway enrichment and immune-stromal TME analysis

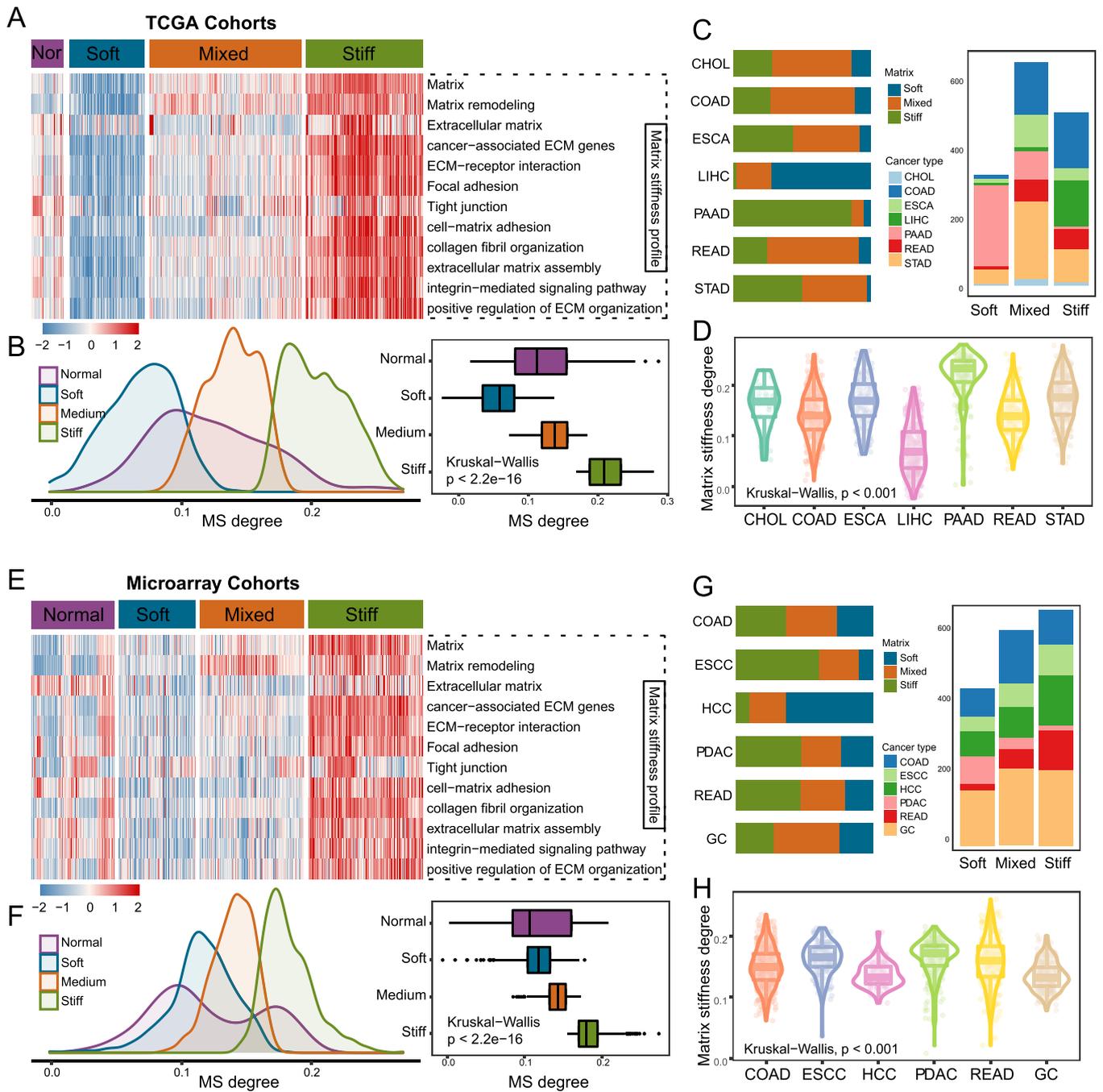
Functional annotation analysis of Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) were performed via the “clusterProfiler” R package. Gene set variation analysis (GSVA) and gene set enrichment analysis (GSEA) were performed to further demonstrate the biological processes based on the Hallmark gene sets, which were downloaded from the MSigDB website (<http://www.gsea-msigdb.org/gsea/index.jsp>). Three algorithms including xCell [28], MCP-counter [28] and EPIC [29] were used to evaluate the immune and stromal cell infiltration of tumors via the “immunedeconv” R package. Publicly available TME infiltration-associated signatures, including “effector cells”, “effector cell traffic”, “CAF signature”, “TAM signature”, “MDSC traffic”, were curated from Bagaev A. et al.’ study [23]. Other TEM-related gene sets, including “antitumor cytokines”, “EMT signature”, “angiogenesis”, “endothelium”, “TGF- $\beta$ -associated ECM”, were also collected from previous studies [23,26]. Stem cell markers and two important stemness indexes (REG-mRNAsi and REG-mDNAsi) were obtained from Malta T. et al. [30]. The above well-established signatures were summarized in Table S2. The activities of TME-related pathways and the expression of representative markers in GI-tumor samples were estimated by ssGSEA algorithm via the “GSVA” R package [31]. In addition, the Stromal Score and Immune Score were calculated using the “ESTIMATE” R package.

### 2.4. Associations of MS subtypes with multi-omics features in pan GI-cancers

Tumor mutational burden (TMB) data was obtained from the “TCGAmutations” R package. MSI (microsatellite instability) sensor scores for TCGA samples are available in the dingMSI data element of the “BiolOncoTK” R package. Somatic copy number variation (level 3, Affymetrix SNP 6.0 array) and somatic mutation (level 4, MAF files) were downloaded from TCGA database (<http://cancergenome.nih.gov/>). Mutations were analyzed and visualized using the “maf-tools” R package.

### 2.5. Construction and validation of the MS-signature by multiple machine learning methods

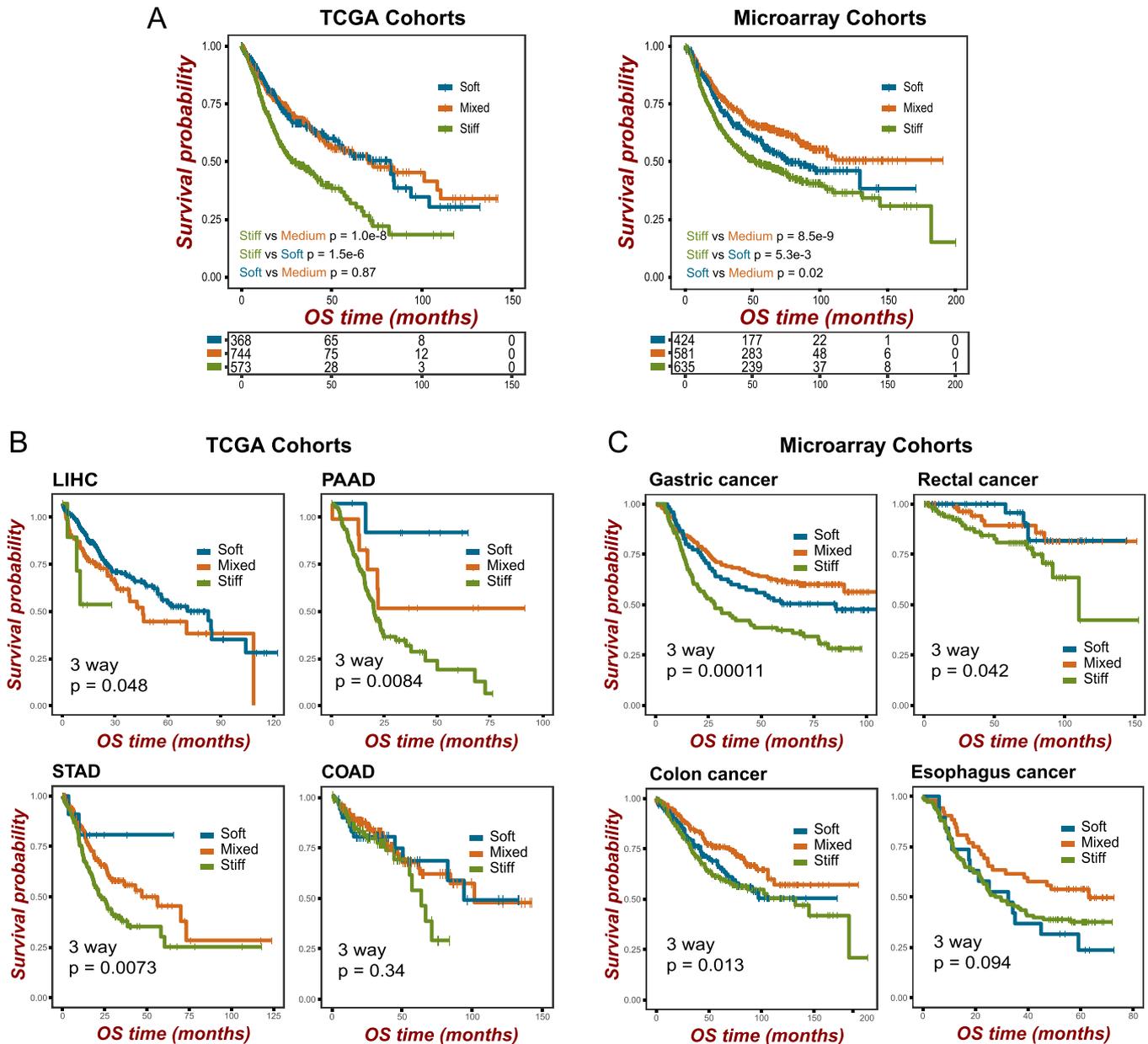
The 568 cancer driver genes were obtained from previous studies [32]. To identify candidate cancer driver genes involved in MS,



**Fig. 1.** Identification of three distinct matrix stiffness (MS)-based subtypes and showing their distribution in GI-cancers. Heat maps shows the MS-specific profiles of the adjacent normal tissues and three subtypes in GI-cancers of the TCGA (A) and Microarray (E) cohorts. High and low ssGSEA scores of 12 MS-specific signatures are represented in red and blue, respectively. Density and box plots showing MS degree among the Soft, Mixed, Stiff subtypes and normal samples of the TCGA (B) and Microarray (F) cohorts. MS degree was the average of 12 ssGSEA scores of MS-specific signatures. Bar charts showing the distribution of three MS subgroups among the different GI-tumor types (left) and the distribution of GI-tumor types among the three MS subgroups (right) in the TCGA (C) and Microarray (G) cohorts. Violin plot showing the MS degree among several GI-cancers of the TCGA (D) and Microarray (H) cohorts. TCGA, The Cancer Genome Atlas; HCC, hepatocellular carcinoma; PDAC, pancreatic ductal adenocarcinoma; GC, gastric cancer.

through the TCGA cohorts, differentially expressed genes (DEGs) between the Stiff and the other tumors were identified using a  $|\log \text{fold change} (\log \text{FC})| > 0.8$  as the threshold of significance. Then univariate Cox regression analysis was used to select the significant genes associated with overall survival (OS) ( $p < 0.0001$ ). Next, the 1685 GI-cancer patients of TCGA cohorts were randomly classified into training ( $N = 1263$ ) and testing ( $N = 422$ ) sets at a ratio of 7.5:2.5. In the training set, extreme gradient boosting (XGBoost), Random Forest (RF), least absolute shrinkage and selection operator (LASSO) regression, and support vector machine (SVM) analyses were performed to select the most important MS-relevant features by

calculating the importance score for each variable via the “XGBoost”, “randomForest”, “glmnet” and “e1071” R packages. The performance of the four machine learning algorithms for feature selection in the training set was evaluated by receiver operating characteristic (ROC) curves, and the areas under curve (AUCs) were subsequently compared. Afterwards, the intersecting genes among the XGBoost, RF, LASSO and SVM analyses were considered as the most pivotal MS-related genes and were visualized by a Venn diagram. Finally, ssGSEA algorithm was performed on the pivotal genes for constructing the predictive model, which was termed by “Matrix Stiffness Score” (MS-score). The predictive performance of the MS-score was



**Fig. 2.** The potential biological pathways and clinical outcomes between three MS subtypes. (A) The association between MS subgroups and OS in GI pan-cancer of the TCGA and Microarray cohorts (Log-rank test). The number of cases and events in a subgroup are shown in the plots. Kaplan-Meier survival plot of overall survival for three MS subgroups in several representative GI-tumor types of the TCGA (B) and Microarray (C) cohorts. TCGA, The Cancer Genome Atlas; OS, overall survival; LIHC, liver hepatocellular carcinoma; PAAD, pancreatic adenocarcinoma; STAD, stomach adenocarcinoma; COAD, colon adenocarcinoma.

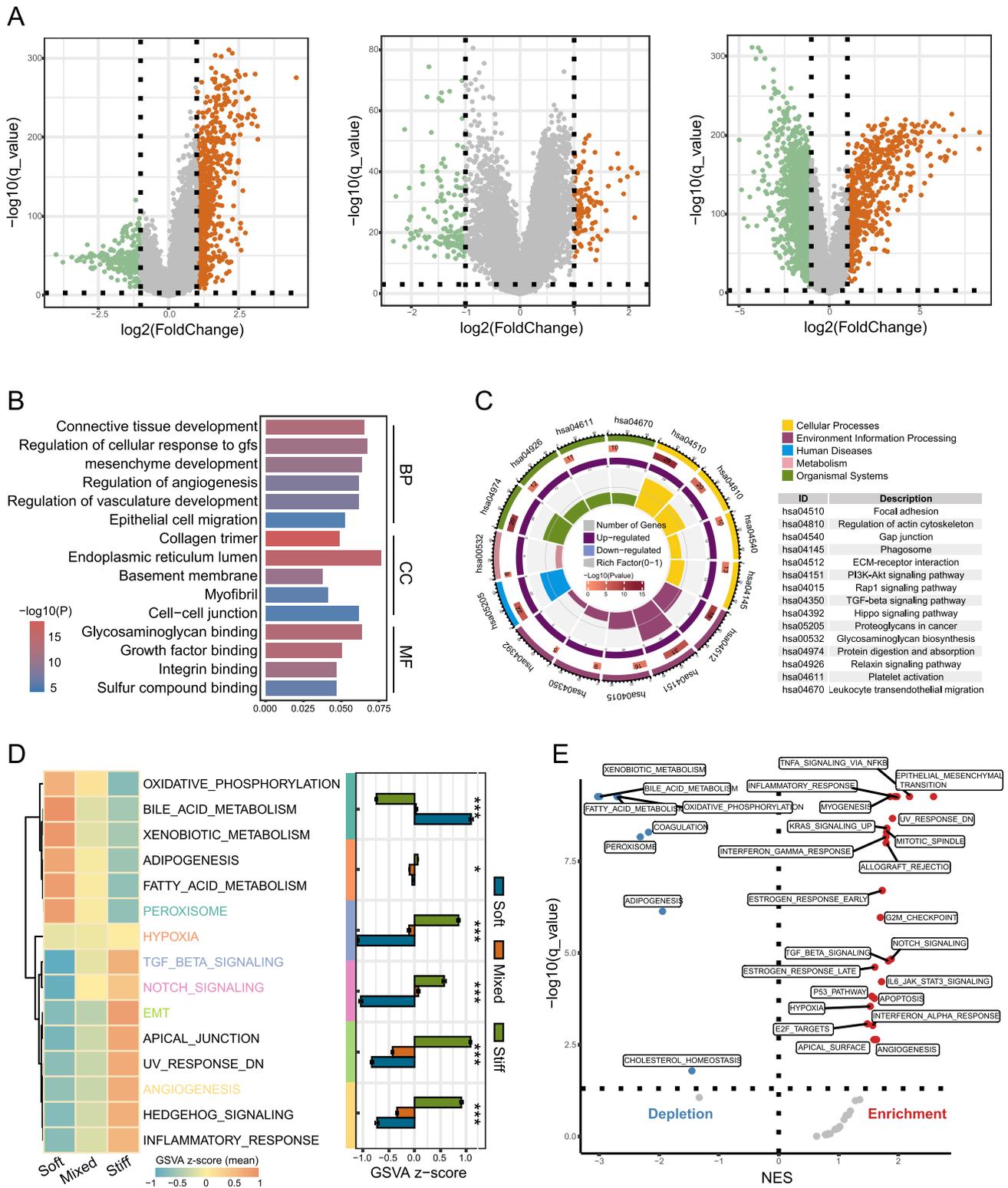
investigated in the training set, test set and external microarray cohorts.

### 2.6. Drug sensitivity analysis

Drug sensitivity and resistance are conventionally quantified by the 50% inhibitory concentration (IC50) [33]. In general, a higher IC50 indicates a greater likelihood of drug resistance. We searched for drug sensitivity data from two authoritative publicly available databases, Genomics of Drug Sensitivity in Cancer (GDSC, <https://www.cancerrxgene.org/>) and Cancer Therapeutics Response Portal (CTRP, <http://portals.broadinstitute.org/ctrp.v2.1/>). The data consisting of IC50 values of compounds in different cell lines and the gene expression profiles of cell lines were organized in the “onco-Predict” R package by Danielle M. et al. [34], and were available for download at <https://osf.io/c6tfx/>. Overall, we obtained 156 (GDSC

and 163 (CTRP) GI-relevant cell lines (Table S3) with both gene expression profile and drug sensitivity data. The MS-scores were calculated based on the gene expression profile of cell lines using the method described above. Information of drug targets was obtained from CTRP (<http://portals.broadinstitute.org/>). The correlation between the MS-score and IC50 values of chemotherapeutic agents, and the relationship between the targets and the MS-signature genes were evaluated by Pearson’s test.

In addition, GSE104580 from the GEO database, a dataset in which HCC patients were receiving transcatheter arterial chemoembolization (TACE), was used as an external microarray cohort for validation and chemotherapeutic sensitivity analysis.



**Fig. 3.** Functional insights of distinct MS heterogeneity between MS subtypes in GI pan-tumors. (A) Three volcano Plots displayed the differentially expressed genes (DEGs) between the Stiff (left), the Mixed (medium), and the Soft (right) versus the rest of tumors, respectively. Top fifteen most significant GO terms (B) and KEGG pathways (C) based on the top 500 DEGs that were screened between the Stiff versus the rest of tumors in the TCGA cohorts. (C) GSEA compared the activation of tumor progressive signaling pathways of Hallmark signatures in three MS subgroups. (D) Volcano plot for enriched (red) and depleted (blue) pathways of Hallmark signatures in the Stiff tumors compared with the other tumors based on GSEA. \*  $P < 0.05$ , \*\*\*  $p < 0.001$ . BP, biological process; CC, cellular component; MF, molecular function; EMT, epithelial–mesenchymal transition.

2.7. Statistical analysis

All statistical analyses were conducted using R version 4.0.1 (<https://www.r-project.org/>). Wilcoxon rank-sum and the  $\chi^2$  test for categorical data were utilized for pairwise comparisons between

groups. The Kruskal–Wallis test was used to compare multiple groups. Correlations between normally distributed variables were assessed with Pearson’s correlation test, while correlations between non-normally distributed variables were assessed with Spearman’s correlation test. The Kaplan–Meier (K–M) curve and log-rank test

were performed to compare survival differences among subgroups using the “survminer” and “survival” R packages. ROC curves and the AUCs were applied to assess the predictive performance using the “pROC” R package. A two-tailed P-value < 0.05 was considered statistically significant.

### 3. Results

#### 3.1. Three GI-tumor subtypes were revealed by unsupervised analysis of the matrix stiffness-specific signatures

To quantify the degree of MS in 1685 bulk tumor samples from 7 GI-tumor types of TCGA, we built a curated list of 12 tumor MS-specific pathways through literature mining (Table S1). The ssGSEA algorithm was used to deconstruct the tumor MS-specific profiles. Then the unsupervised clustering analysis based on the MS-specific profiles of GI-tumors can be clustered into three distinct matrixes (Fig. S2). According to their activation in the tumor MS-specific pathways, we termed these three clusters as “Soft”, “Mixed”, and “Stiff” (with 368, 743 and 574 cases, respectively; Fig. 1A). Interestingly, we observed the 12 tumor MS-specific pathways in normal samples was more silent than the Stiff but similar to the Soft and Mixed subtypes. Here, the average of 12 ssGSEA scores of MS-specific signatures was referred to as “MS degree”, roughly representing the overall stiffness of each sample’s matrix. As expected, the MS degree in the Soft, Mixed and Stiff subtypes of GI-tumors increased hierarchically (Fig. 1B). The MS degree of normal tissues was between the Soft and Mixed subtypes. Similar results were found in the Microarray cohorts (Fig. 1E, F). In brief, the MS heterogeneity existed in GI-cancers.

As the bar charts shown (Fig. 1C, G), different GI-tumor types varied substantially in their composition of MS subgroups. The Soft subgroup was particularly enriched in LIHC and the Stiff subgroup was obviously enriched in PAAD. Furthermore, by comparing the MS degree of different GI-tumors, we found that PAAD was the highest one (Fig. 1D, H), indicating significant alteration in the tumor matrix of PAAD.

#### 3.2. Distinct GI-cancer prognoses existed in three MS subtypes

K-M survival analysis revealed that patients in the Stiff subtype of GI-cancer had significantly shorter overall and disease-specific survival times than those in the Soft and Mixed subtypes. The Mixed subgroup was associated with a superior progression-free survival in the TCGA cohorts and OS in the Microarray cohorts (Fig. 2A, Fig. S3A). Additionally, among different types of GI-cancers, three MS subtypes showed distinct survival outcomes both in the TCGA (Fig. 2B, Fig. S3B) and Microarray cohorts (Fig. 2C, Fig. S3C). Regarding to the relationship between the MS subtypes and clinical features, there were larger percentages of Stage I or Grade 1 tumors in the Soft subtype than the Mixed and Stiff tumors (Fig. S4). Taken together, patients in the Stiff subtype of GI-cancers were closely associated with poorer clinical features and prognosis.

#### 3.3. Distinct functional and biological pathways were observed in three MS subtypes

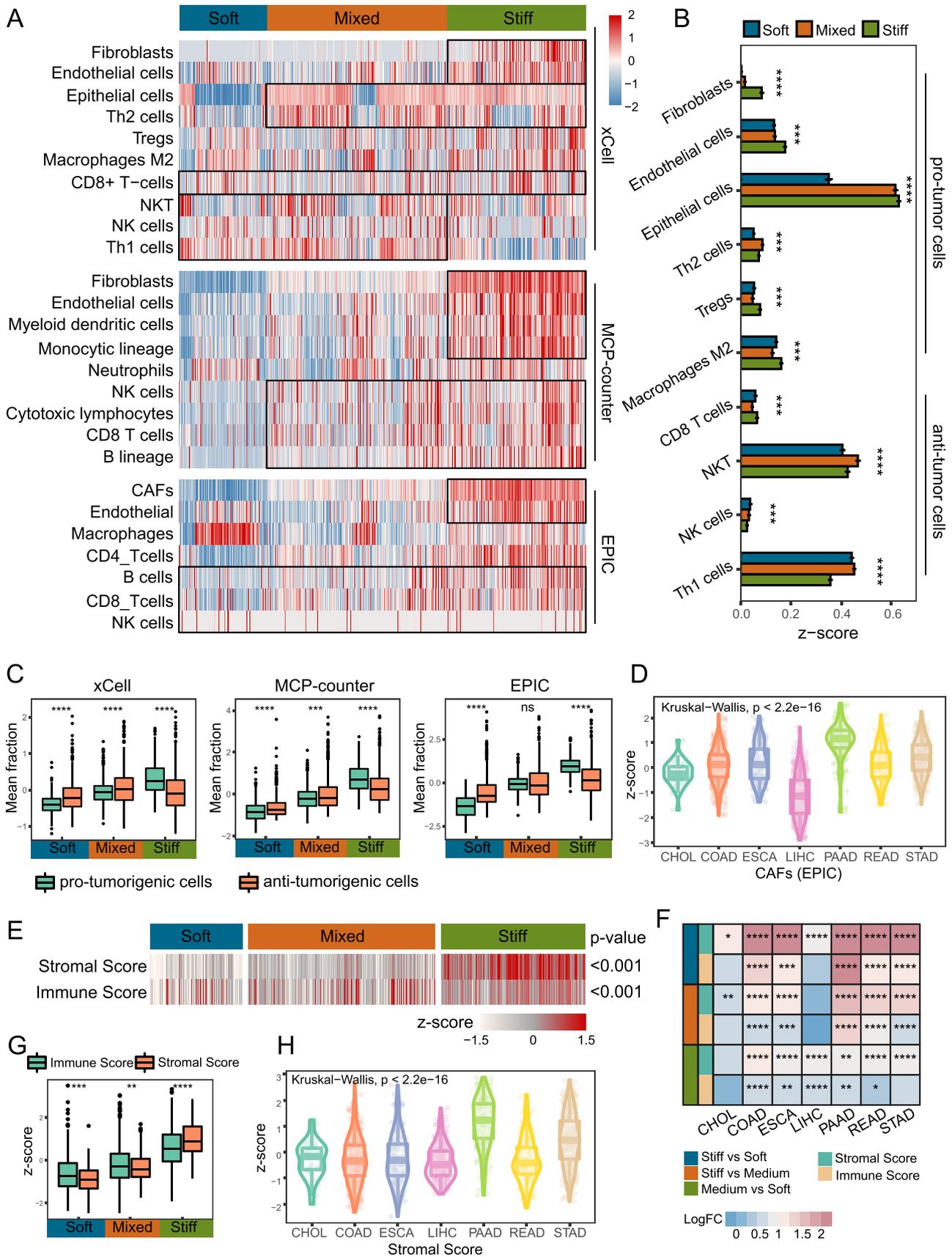
Differential expression analyses were performed in the Stiff, Mixed, and Soft subtypes sequentially (Fig. 3A). Significantly different gene expression profiles existed between the Soft subtype and the others and between the Stiff tumors and the others. To investigate potential biological behaviors between MS subgroups, functional enrichment analyses were applied. Separated GO and KEGG analyses for the upregulated DEGs in the Soft and the down-regulated DEGs in the Stiff tumors showed similar results. They were significantly enriched in functions and pathways related to collagen

and ECM, such as collagen trimer, collagen binding, ECM structural constituent, ECM–receptor interaction, and focal adhesion (Figs. 3B–C and S5A–B). In addition, GSVA on Hallmark features inferred that some tumor progression-related pathways, such as TGF- $\beta$  signaling, notch signaling, epithelial–mesenchymal transition (EMT) and angiogenesis, were highly enriched in the Stiff tumors but silent in the Soft subtype (Fig. 3D). The subsequent GSEA further confirmed the conclusion (Fig. 3E). Hallmark pathways varied substantially not only among the MS subgroups, but also among the GI-tumor types (Fig. S6). Of note, PAAD had the highest GSVA scores in the EMT, TGF- $\beta$  signaling and angiogenesis pathways. Whereas LIHC had the lowest scores in the TGF- $\beta$  signaling pathway. Tumors in the bottom of the enrichment of EMT and angiogenesis included COAD, READ and LIHC. On the contrary, they had relatively high scores of the peroxisome signal. These results indicated that PAAD may have the most significant MS alteration, be associated with the stiffest tumor matrix, and exhibit obvious malignancy among the GI-cancers.

#### 3.4. MS-based GI-cancer subtypes showed distinct immune and stromal microenvironments

To explore the TME cell infiltration in each of the three MS subtypes of GI-cancer, three algorithms including xCell, MCP-counter and EPIC were applied (Fig. 4A). According to the xCell algorithm, the fibroblasts and endothelial cells infiltrated significantly in the Stiff tumors, and the infiltration of epithelial cells and Th2 cells was more enriched in the Stiff and Mixed subtypes than the Soft. The above cells were known to be associated with tumor progression [35]. Regarding the infiltrations of the anti-tumor cells such as CD8 + T cell, natural killer (NK) T cell and Th1 cell, they showed no significant superiority in the Stiff tumors compared with the others. The bar chart displayed more intuitively the differences of immune infiltration between MS subtypes (Fig. 4B). Consistent with the xCell data, results based on the MCP-counter and EPIC algorithm showed the pro-tumor cells such as fibroblasts, endothelial cells, myeloid dendritic cells, monocytes, and especially cancer-associated fibroblasts (CAFs), exhibited a state from depleted to enriched with the hierarchical tumor matrix from the Soft to the Stiff. Interestingly, the anti-tumor cells also displayed a slight increase from the Soft to the Stiff tumors. Next, we investigated which pro-tumor or anti-tumor cells were dominant in different MS subtypes (Fig. 4C). Notably, the enrichment of pro-tumor cells in the Stiff tumors was significantly higher than that of anti-tumor cells according to all three algorithms. On the contrary, although the enrichment levels of all immune and stromal cells in the Soft tumors were inferior to the Stiff, the anti-tumor cells rather than the pro-tumor cells were dominant in the Soft and Mixed subtypes. CAFs participate in the synthesis of ECM components such as collagen and fibronectin, contributing to tumor matrix stiffening [36]. Expectedly, PAAD had the highest degree of CAFs infiltration among the GI-cancers (Fig. 4D).

Then, ESTIMATE analysis was applied to further demonstrated the stromal and immune features among the MS subtypes of GI-cancer. Heat maps firstly showed the distribution of Immune Score and Stromal Score in three MS subtypes, with both scores showing hierarchical increases the Soft to the Stiff group (Fig. 4E). Then the comparative analysis revealed that the Stromal Score had more increment than the Immune Score with the MS degree increasing across seven GI-cancer types (Fig. 4F). Finally, the box diagram displayed that the Immune Score were higher than the Stromal Score in the Soft and Mixed tumors, which was the opposite in the Stiff tumors, indicating that the stroma activation was superior to the immune response in the Stiff tumors (Fig. 4G). Unsurprisingly, the Stromal Score of PAAD was the highest across seven GI-cancer types (Fig. 4H).



(caption on next page)

**Fig. 4.** Distinct immune and stroma microenvironments in the MS-based GI-cancer subgroups. (A) Heat map showing the profiles of infiltrating immune cells based on the xCell, MCP-counter and EPIC algorithm in the three MS subgroups. (B) Bar charts showing the differences of the immune cell infiltration according to the xCell algorithm among the three MS subgroups. (C) Box plots showing the comparison in the fraction between the pro-tumor and anti-tumor cells within the three MS subgroups according to three algorithms. In the xCell data, cells that may promote immunosuppression were identified as pro-tumorigenic cells, including macrophages M2, Tregs, Th2, epithelial cells, endothelial cells, and fibroblasts; the anti-tumorigenic cells involved in killing tumor include CD8 T, NKT, NK, and Th1 cells. In the MCP-counter data, the potential pro-tumorigenic cells include fibroblasts, endothelial cells, myeloid dendritic cells, and monocytic lineage; the anti-tumor cells include NK cells, cytotoxic lymphocytes, CD8 T cells, and B lineage. In the EPIC data, the potential pro-tumor cells include CAFs and endothelial cells; the anti-tumor cells include B cells, CD8 T cells, and NK cells. (D) Violin plot showing the infiltration score of CAFs among several GI-cancers. (E) Heat map showing Stromal Score and Immune Score based on the ESTIMATE algorithm in the three MS subgroups. (F) The comparison analysis of the Stromal Score and Immune Score among MS subtypes. (G) Box plots showing the comparison between the Stromal Score and Immune Score within the three MS subgroups. (H) Violin plot showing the Stromal Score among several GI-cancers. \* $P < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , \*\*\*\* $p < 0.0001$ . CAFs, cancer-associated fibroblasts.

The above results including enrichment analysis and immune-stromal infiltration have roughly clarified the distinct TMEs across MS subtypes of GI-tumors. We subsequently verified the conclusions by ssGSEA algorithm based on the published specific signatures (Fig. 5A). Of note, there was no significant differences in the enrichment scores of antitumor cytokines across the three MS subgroups, while the TGF- $\beta$ -associated ECM, EMT and angiogenesis signals were significantly activated in the Stiff subgroup, indicating the malignant and proliferative characteristics in the Stiff tumor (Fig. 5B). The stem cell markers were markedly activated in the Mixed and Stiff subtypes of GI-cancers (Fig. 5A). Previous study reported that increased stemness was associated with higher tumor invasiveness and worse clinical features [30]. In our study, two important stemness indices proposed by Malta T. et al. [30], the epigenetically regulated mRNA expression-based and DNA methylation-based stemness indexes (REG-mRNasi and REG-mDNasi) were compared between the MS groups (Fig. 57), in which the REG-mRNasi was higher in the Mixed and the Stiff groups than the Soft tumors. Compared with the antitumor cytokines, the ssGSEA scores of TGF- $\beta$ -associated ECM, EMT and angiogenesis were higher in the Stiff groups and lower in the Soft and Mixed groups (Fig. 5C), suggesting that pro-tumor factors in the Stiff tumors were stronger than anti-tumor immunity, leading to its more malignant features. Furthermore, there were significantly positive correlations between pro-tumor factors (such as CAF, MDSC traffic, etc.) and the MS degree, but relatively weak correlations were shown between the anti-tumor immunity (including antitumor cytokines and effector cells) and the MS degree (Fig. 5D), suggesting that the malignant biological behaviors were activated with the tumor stiffening. Collectively, the Stiff tumor was characterized by enriched infiltration of pro-tumor cells, especially CAFs, and activated states of tumor progression-related pathways (Fig. 5E). These events may form a complex communication network with each other, participating in shaping a “stiff” tumor matrix to hold tumor malignancy.

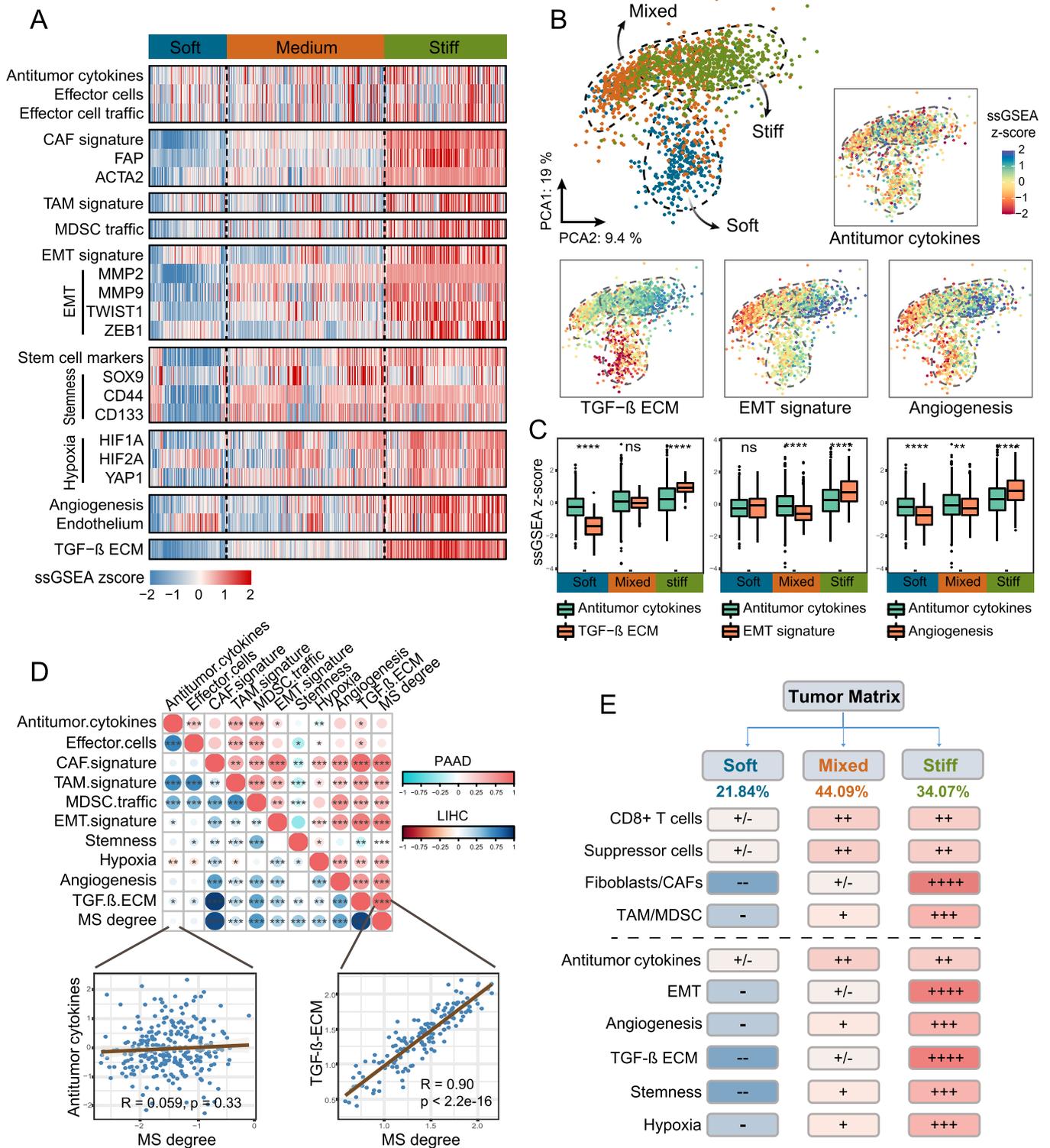
### 3.5. Multi-omics alterations in the GI-cancer subgroups according to MS classification

Next, we investigated frequency differences of somatic and copy number variation (CNV) mutations across the MS subtypes (Fig. 6A). Specifically, the Stiff subtype had a significantly highest frequency of TP53 and KRAS (56% and 31%, respectively). The Mixed subtype had higher TMB, higher MSI score, higher genome deletion and amplification rates among the MS subgroups of GI-tumors (Fig. 6B). The deletion and amplification regions on chromosomes also varied from different MS subtypes (Fig. 6C). Moreover, we compared the MS subtypes to the existing molecular subtypes (Fig. 6D). We found that the Stiff subgroup had a high overlap with the EMT (STAD), desmoplastic (PAAD), stroma activated (PAAD) and diffuse (STAD) subtypes, while the Soft tumor had high proportion overlapping with the MSI (READ), TMB-upper (ESCA) and immune classical (PAAD) subtypes, which were consistent with our conclusions earlier.

### 3.6. Machine learning revealed a MS-signature for predicting the Stiff tumor, and validated the MS classification in two independent cohorts

Given that the stiffened matrix was potentially associated with tumor progression, we next probed the expression profiles of 568 cancer driver genes in the MS subgroups, which identified 58 cancer driver genes closely associated with the Stiff tumors ( $|\log\text{FC}| > 0.8$ ) (Fig. 7A, Table S4). Among them, 42 of the 58 MS-related cancer driver genes affected significantly the survival prognosis based on the univariate Cox analysis ( $p < 0.0001$ , Table S4). The GI pan-cancer data was randomly compartmentalized into the training and test cohorts using a ratio of 7.5:2.5. Subsequently, we applied XGBoost, RF, LASSO and SVM algorithms, and 29, 20, 33, and 29 genes critical for the Stiff matrix were identified respectively. The ROC curve analysis demonstrated that the four machine learning algorithms had good performance in feature selection, with AUCs  $> 0.90$  in the training set and AUCs  $> 0.85$  in the testing set, indicating that a good model can be built to predict the Stiff and the other tumors (Fig. 7B). Combining the results of four machine learning methods, 11 common genes were identified as the most pivotal MS-related cancer driver genes, and formed a MS-signature (Fig. 7C). A protein-protein interaction network for the 11 pivotal MS-related diver genes was displayed (Fig. 7D). Expression levels of these MS-related cancer driver genes increased from the Soft to Stiff tumors, and provided additional prognostic value in the individual MS subgroups. They were found to be significantly associated with the activation of CAF, EMT, angiogenesis and TGF- $\beta$ -associated ECM pathways but not with the antitumor cytokines (Fig. 7E). Furthermore, based on the 11-gene MS-signature, Matrix Stiffness Score (MS-score) was constructed by ssGSEA algorithm to quantify the MS of each tumor sample. The optimal cutoff value for discrimination was 0.562, suggesting that patients with a score  $> 0.562$  were considered to be the Stiff subtype and those with a score  $< 0.562$  were considered to be the others (Mixed and Soft). Strikingly, the MS-score had an excellent performance in distinguishing the MS subtypes as evaluated in the test set, with an AUC of 0.980, a sensitivity of 86.16%, a specificity of 97.72%, and an accuracy of 94.79% (Fig. 7F). Higher MS-score referred remarkably worse survival outcomes both in the TCGA and Microarray cohorts (Fig. 7G), indicating that the MS-score can serve as a biomarker for GI-cancer prognosis. Importantly, we proved the MS-score was remarkably positively correlated with the tumor progression-related pathways, but weakly correlated with the anti-tumor signals (Fig. 7H), which was also consistent with earlier conclusions.

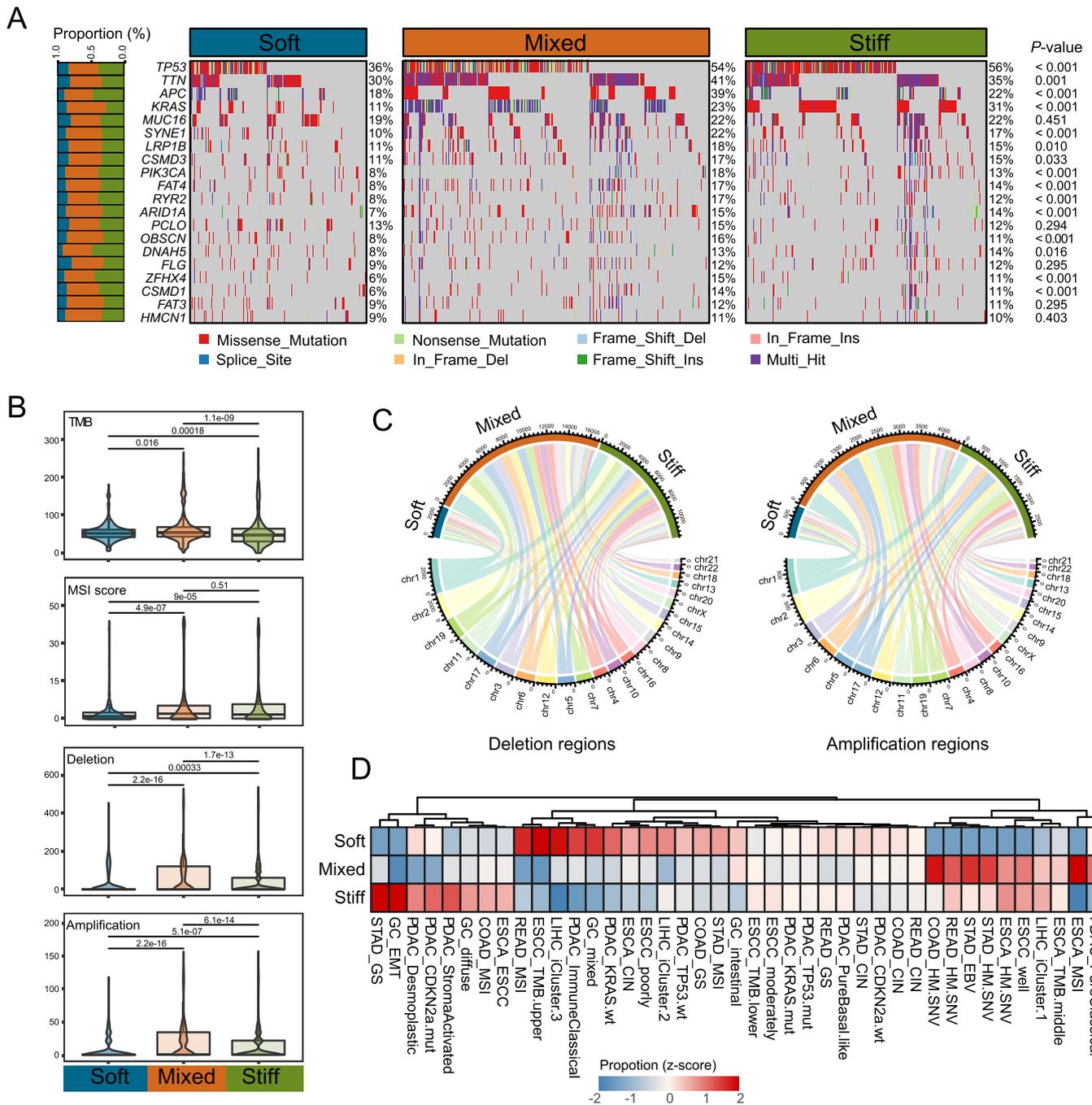
We verified the performance of the MS-score in two independent cohorts (PAAD and HCC). Using the MS-score partitioning, more than three-quarters (88.8%) of the samples in the PAAD cohort were identified as the Stiff tumors, compared to only 36.7% of the HCC cohorts (Fig. 7I, L). These stiff tumors were characterized by being activated significantly in MS-specific pathways as earlier. Also, the EMT, angiogenesis and TGF- $\beta$ -associated signals of the Stiff were higher than the others, while the antitumor cytokines showed no significant difference between the Stiff and the others (Fig. 7J, M). In the PAAD cohort, patients with higher MS-score had both worse OS and disease free survival outcomes, further verified the tumor with



**Fig. 5.** MS-based GI-cancer subtypes showed distinct TMEs. (A) Heat map showing the activation of biological processes and typical molecules among MS subtypes of GI-cancer. The ssGSEA algorithm was used to quantify the activity of TME-related signaling pathways and the expression of representative markers. High and low ssGSEA z-scores were exhibited in red and blue, respectively. (B) PCA plots showed the correlation between the activation of several biological pathway and MS subtypes. (C) Box plots showing the comparison between the activation of biological pathways within the three MS subgroups. (D) Correlation between the biological process activation and MS degree in LIHC and PAAD. (E) Schematic description of the features associated with the three GI-cancer MS subtypes. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , \*\*\*\* $p < 0.0001$ . EMT, epithelial-mesenchymal transition; PCA, principal components analysis; CAFs, cancer-associated fibroblasts; ECM, extracellular matrix; TAM, tumor associated macrophage; MDSC, myeloid-derived suppressor cell.

stiffer matrix was associated with worse prognosis (Fig. 7K). In a HCC cohort (GSE104580) in which all patients received TACE treatment, the MS-score was significantly lower in the response group, and the response rates were obviously lower in the Stiff tumors than

in the others (Fig. 7N), which illustrated that high MS-score and the stiffened tumor matrix were related to chemotherapy resistance. These data were consistent with the conclusions of our analysis above, demonstrating the excellent performance of the MS-score.



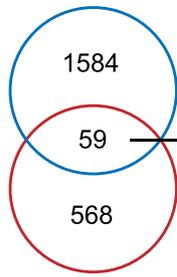
**Fig. 6.** Correlation of multi-omics alterations and drug sensitivity with MS-based GI-cancer subgroups. (A) Landscape of mutation status of top 20 frequently mutated genes among the MS subtypes. The proportion of mutation rates of these genes in the MS subtypes are denoted on the left. (B) Violin plots displaying the differences in tumor mutation burden, MSI sensor scores, genome deletion and amplification among the three MS subgroups. (C) Circular plot showing the deletion and amplification regions in the three subtypes. (D) Overlay of the three MS subgroups with existing TCGA molecular subgroups. Each row indicated the distribution of MS subgroups within each molecular subtype. Red revealed a higher proportion, whereas blue revealed a lower proportion. \* $P < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , \*\*\*\* $p < 0.0001$ .

### 3.7. Efficacy of the MS-score in predicting drug sensitivity

In the above analyses, we clarified that the malignant behaviors such EMT, angiogenesis and TGF- $\beta$ -associated signals in tumors would be hierarchically activated from the Soft to Stiff tumor matrix. Previous literature reported that EMT pathway and TGF- $\beta$ -associated ECM in the tumor were closely related to chemotherapy resistance [37]. In our study, we explored the relationship between the MS-signature and drug sensitivity using the pharmacological data from GDSC and CTRP databases. The landscape of the correlation between drug sensitivities and the MS-score are showed in Fig. 8A. We found

that the IC50 of many chemotherapy drugs commonly used in cancers such as fluorouracil, trimetatinib, oxaliplatin and crizotinib was positively correlated with the MS-score (Fig. 8B). The relationships of the IC50 values of these commonly used drugs and the MS-score in individual cancer types of cell lines were shown in Fig. S8. The correlation between MS-signature genes and classical therapeutic targets in cancer was showed in Fig. 8C. These results suggest that patients with high MS-score were resistant to standard chemotherapy regimens, and the MS-score could serve as a promising indicator to predict chemotherapy resistance.

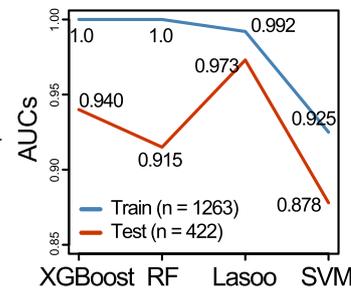
**A Stiff vs others DEGs**



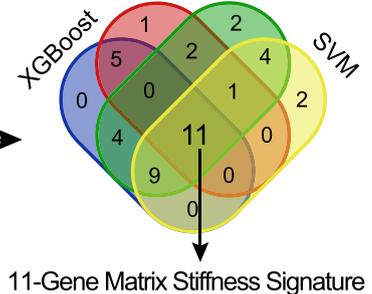
Univariate cox  
(n = 43)

Cancer driver genes

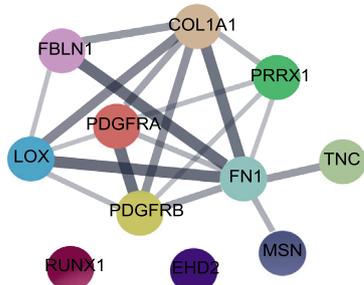
**B**



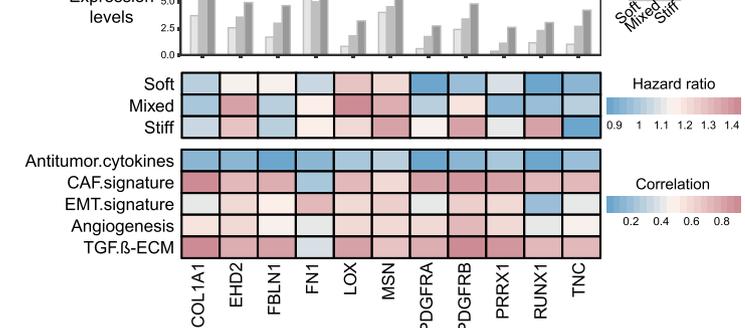
**C**



**D**



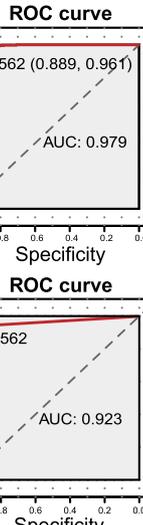
**E**



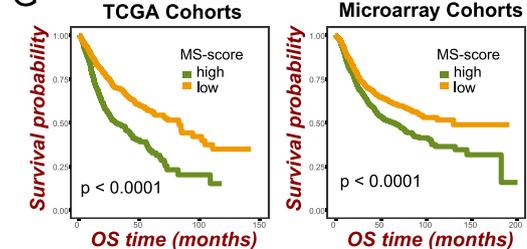
**F**

**Matrix Stiffness Score**

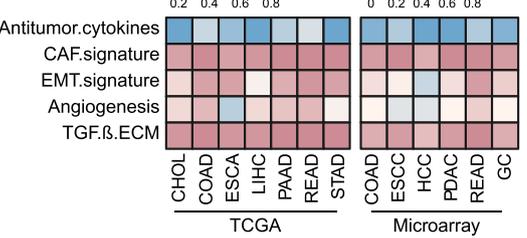
	Train set (N = 1263)		Test set (N = 422)	
	Predict Stiff Tumors	Predict Other Tumors	Predict Stiff Tumors	Predict Other Tumors
Predict Stiff Tumors	414	92	137	6
Predict Other Tumors	17	740	22	257
Totals	431	832	137	263
Correct	414	740	159	257
Sensitivity (%)	96.06		86.16	
Specificity (%)		88.94		97.72
Accuracy (%)	91.37		94.79	



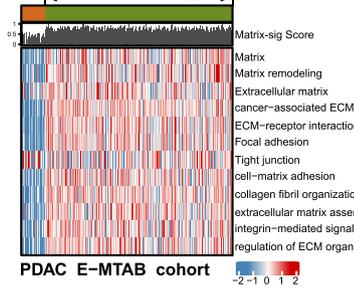
**G**



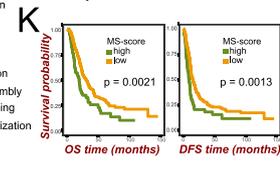
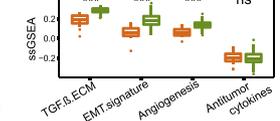
**H**



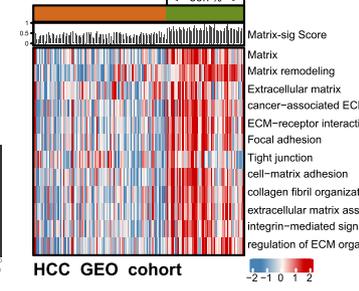
**I**



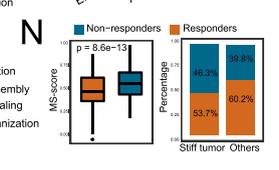
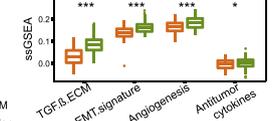
**J**



**L**



**M**



(caption on next page)

**Fig. 7.** Construction and validation of the MS-score. (A) Venn diagram of intersection of DEGs and cancer driver genes in the GI-cancer samples of TCGA cohort. The cut-off criteria of DEGs were  $|\log_{2}FC| > 0.8$  and an adjusted  $p < 0.05$ . The intersected genes were further screened by Univariate Cox analysis ( $p < 0.0001$ ). (B) The performances of four machine-learning algorithms for feature selection were, respectively, evaluated in the train set and test set. AUCs were generated by ROC analysis. (C) Venn diagram identified 11 of the most critical MS-specific genes that were shared by four feature selection algorithms. (D) The protein-protein interaction network is formed using the key MS-associated driver genes. (E) Gene-level summary of the genomics features in the GI pan-cancer samples. Bar plot shows the expression of 11 crucial MS-associated driver genes. Heat maps show the hazard ratios of key MS-associated driver genes, the correlation between the expression of these MS-related driver genes and the activation of several typical biological processes in the three MS subgroups. (F) Left panel: confusion matrices of binary results of the MS-score for the train set (upper) and test set (lower). Right panel: ROC curves of the MS-score in distinguishing the Stiff tumor and the other subtype in the train set (upper) and test set (lower). (G) Kaplan–Meier survival analysis of the MS-score in the TCGA (left) and Microarray (right) cohorts. (H) Correlation of the MS-score with the typical biological processes. Validation of the performance of the MS-score in distinguish the Stiff subtype and the other tumors via two independent cohorts, a PDAC cohort (I) and a HCC cohort (L). Box plots displayed the activation of the typical biological processes among the MS subtypes in the independent PDAC (J) and HCC cohort (M). (K) Kaplan–Meier survival analysis showed that higher MS-score referred worse OS and PFS for the PDAC patients. (N) The MS-score was significantly higher in the response group (left), and the response rates were obviously higher in the Stiff tumors than in the others (right). \* $P < 0.05$ , \*\*\* $p < 0.001$ . XGBoost, extreme gradient boosting; RF, Random Forest; LASSO, least absolute shrinkage and selection operator; SVM, support vector machine; ROC, receiver operating characteristic; AUC, areas under curve; TCGA, The Cancer Genome Atlas; OS, overall survival; DFS, disease-free survival; HCC, hepatocellular carcinoma.

#### 4. Discussion

The emergence of genomic analysis of GI-cancers enabled us to recognize the potential biological and physical factors that shape a tumor. Based on these analyses, we can identify predictive and prognostic subgroups, and provide new insights into clinical diagnosis and treatment for GI-cancers [4]. Matrix stiffness (MS) is recognized as a critical factor in cancer progression [16]. Although measurement methods of the MS boom out nowadays, their detections are only limited in physiological level and they are not suitable for all types of GI-tumors [38]. It is urgent to distinguish a “soft” or “stiff” tumor from the level of matrix molecules, and explore the biological role of MS in driving tumor growth. Herein, we comprehensively characterized the landscape of MS subtypes identified using literature-derived matrix-specific signatures within seven GI-cancer types of TCGA. Three predicted subtypes were validated in the Microarray cohorts, clearly demonstrating the robustness of the MS classification. Distinct biological processes, immune and stromal infiltration, and mutation features between the MS subtypes were further revealed. Of note, the subtypes we developed overlapped to a certain extent with the existing molecular subtypes of GI-cancers, confirming the credibility and availability of the MS subtypes in GI-cancers. To facilitate the distinction of the MS subtypes in clinical practice, eleven most critical MS-related genes were recognized and defined as the MS signature by using the XGBoost, RF, LASSO and SVM machine learning methods, and its robustness were subsequently demonstrated.

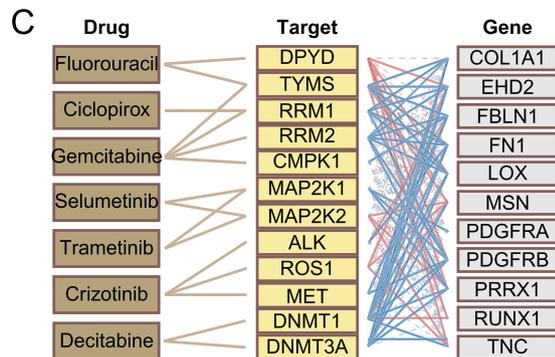
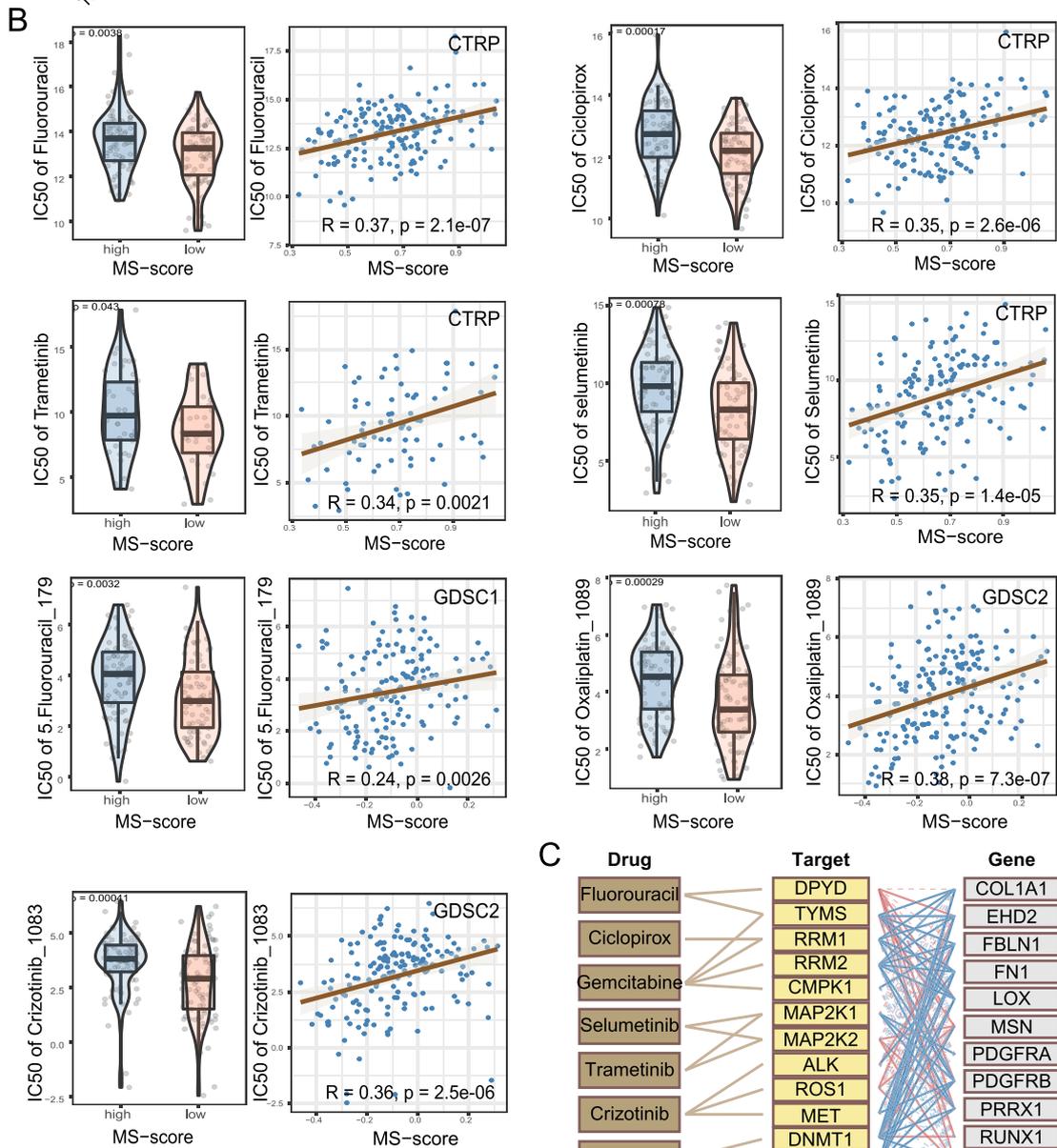
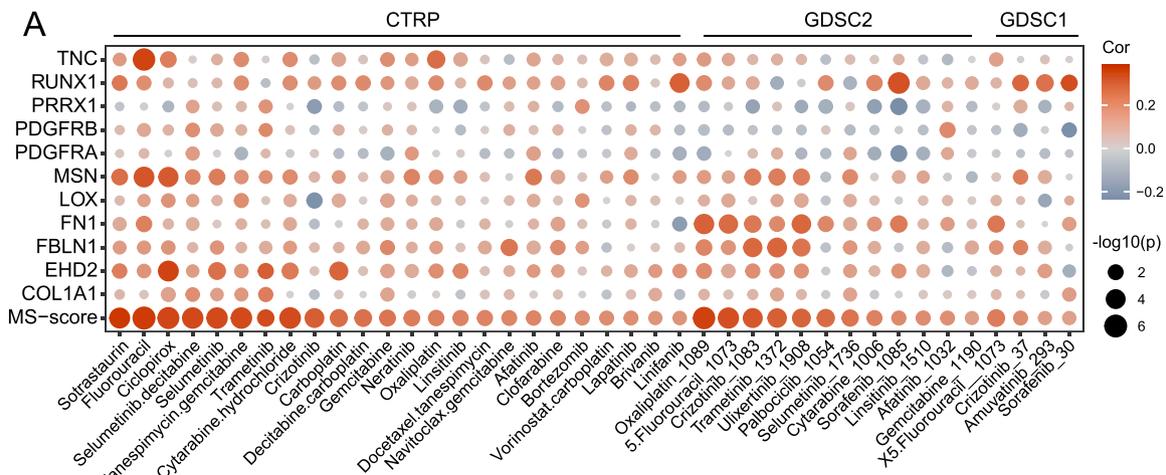
In solid cancers such as breast and pancreatic cancer, tumors often contain abnormally stiff tissues, mainly caused by the accumulation, contraction, and crosslinking of the ECM [16]. However, evident heterogeneity exists between tumors from different patients within one type cancer [39]. In this study, based on the MS-specific signatures, GI-tumors were classified into three hierarchical MS subtypes, termed as “Soft”, “Mixed”, and “Stiff”. Through comparing the MS degree among various GI-cancer types, we found the PAAD had the highest level, suggesting the most significant stiffening alteration of the matrix. In effect, the TME of PAAD was characterized by a prominent desmoplastic reaction, accompanied by activated CAFs and substantial deposition ECM, which accounted for up to 90% of the tumor [40,41]. In addition to the MS as the tumor internal force, solid stress exerted by the surrounding tissue is an external force during tumorigenesis [12]. As is well-known, a large proportion of HCC develop from liver fibrosis or cirrhosis [42], so that the tumor harbored large solid stress from cirrhotic tissue, whereas the alteration at the level of matrix molecules will matter less. To some extent, this can explain why the majority of liver cancer samples were “soft” in this study.

It is now increasingly accepted that the TME is a heterogeneous collection of infiltrates, in which stromal and immune cells are important components [8]. The immune and stromal infiltration plays both anti-tumor and pro-tumor roles in tumor progression [35]. Among the MS subtypes of GI-tumors in our study, the infiltration of pro-tumor cells, especially fibroblasts and CAFs, were increased

hierarchically from the Soft to Stiff subtype. CAFs are major contributors to an immunosuppressive TME by producing immunosuppressive cytokine TGF- $\beta$  in the cancer stroma [43,44]. In addition, CAFs can remodel the ECM by secreting collagen and fibronectin to enhance the stiffness of tumor [45,46], which impede trafficking of T cells (especially CD8+T cells), thereby suppressing anti-tumor immunity [47]. Taken together, a “stiff” tumor matrix infiltrated abundant CAFs may shape an immunosuppressive tumor stromal microenvironment. Of note, by the comparison within MS subgroups, we found the Stiff tumors showed higher pro-tumor cell infiltrating fraction than the anti-tumor cells, while the Soft and Mixed tumors were on the opposite. This data indicated that the pro-tumor response was stronger than anti-tumor immunity in the Stiff tumor, leading to more malignancy and worse prognosis of the Stiff tumor. Instead, the Soft and Mixed tumor showed relatively favorable prognosis.

Actually, ECM stiffening does not operate the cancer cell invasion in isolation, but rather forming a complex communication network with other molecules within the TME. ECM stiffness activates epidermal growth factor receptor in cancer cells, then induce EMT promoting cancer invasiveness [48]. The invasion independent of EMT is also mediated by TGF- $\beta$  signaling during ECM remodeling [49]. In addition, under stimulation of TGF- $\beta$ , fibroblasts are recruited to the site of tumor and transformed to CAFs, further enhancing the MS and pursuing tumorigenic features [50]. CAFs have interactions with various cells, including endothelial cells that promote angiogenesis [36]. Consistently, the Stiff subtype of GI-cancer in our study was characterized by significant activated pathways of TGF- $\beta$ , EMT and angiogenesis, but also activated molecules of stemness. Indeed, a stiff matrix inducing stemness of cancer cells in melanoma [51], HCC [52] and glioma [53] has been reported. By contrast, these pro-tumor signals were silent in the Soft tumors. In contrast to these malignant pro-tumor pathways, no significant difference was found in the activation of anti-tumor cytokines across three MS subtypes. Therefore, in the Stiff subgroup, pro-tumorigenic factors were dominant rather than anti-tumorigenic factors, consistent with the immune and stromal infiltration of pro- and anti-tumorigenic cells. This further confirmed the malignancy of the Stiff subtype and explain the unfavorable prognosis of this subtype GI-tumor. Taken together, such molecules and cells were interconnected within TME, contributing to shaping a “stiff” tumor and immunosuppressive TME, thereby promoting cancer progression.

Despite advances in chemotherapy and molecularly targeted drugs, drug insensitivity and chemoresistance are two tough challenges for management of GI-cancers. MS has been revealed to promote drug resistance in mammary cancer cells [54], HCC cells [55] and pancreatic cancer cells [56], but soft matrix has also been reported to induce drug resistance in laryngeal squamous cell carcinoma cells [57]. In our study, the MS-score constructed by machine-learning methods was positively correlated with the IC50 of multiple drugs in GI-relevant cells, indicating that the sensitivity of chemotherapy decreases with the tumor matrix stiffening. ECM stiffness inducing chemoresistance may be owing to the matrix



(caption on next page)

**Fig. 8.** Efficacy of MS-signature in predicting drug sensitivity. (A) Bubble plot showing the relationship between the IC50 values of commonly used chemotherapy drugs, the MS-score, and the MS-signature genes through the data of GI-cancer related cell lines from CTRP and GDSC databases. The IC50 data of compounds in different cell lines and the gene expression profiles of cell lines were downloaded from <https://osf.io/c6tfx/>. The MS-score was calculated by the method described earlier based on the gene expression profiles in the CTRP and GDSC databases. (B) Boxplots of the comparison of IC50 of drugs between high- and low-MS-score groups, and correlation between the IC50 and MS-score. (C) The correlation between the MS-signature genes and classical therapeutic targets commonly used chemotherapy drugs. The data of drug targets was obtained from the CTRP website (<http://portals.broadinstitute.org/>). The red line represents a positive correlation and the blue line represents a negative correlation. CTRP, Cancer Therapeutics Response Portal; GDSC, Genomics of Drug Sensitivity in Cancer.

structure with high deposition of collagen and proteoglycans acts as a dense barrier for prevention of drug absorption or delivery [15]. In addition, the Stiff subtype exhibited significant activation of EMT program and TGF- $\beta$  signaling, which are two critical contributors to anti-cancer drug resistance [58,59]. Thus chemotherapy sensitivity decreasing in the Stiff tumor can be explained. Moreover, most of the 11 MS-signature genes were negatively correlated with many classical cancer therapeutic targets, further suggesting the MS-signature could be biomarkers for predicting the development of chemotherapy resistance. Anti-stromal therapies are showing considerable promise in the treatment of some tumors, such as PDAC [60], a GI-cancer type that has higher MS degree in our study. Stromal inhibitors may have the potential to improve chemotherapy sensitivity in the patients with high MS-score, but further in vitro and in vivo studies are needed to classified the specific effects and mechanisms.

## 5. Conclusion

In summary, our study revealed the MS heterogeneity in GI-cancers and advocate to term the MS subtypes as “Stiff”, “Mixed”, and “Soft” due to their distinct molecular features in the tumor matrix. We described the differences of biological behaviors between MS subtypes, affording the biological interpretability for the difference of prognosis. Besides, the genomic and transcriptional landscape of MS subtypes were preliminarily revealed. Moreover, to provide predictions for MS subtypes and chemotherapy response, a robust and concise model including only 11 pivotal MS-related genes was established. The comprehensive characterization of the hierarchical MS subtypes within the TME facilitates our understanding of the MS heterogeneity in GI-cancers, helps the clinical stratification and assists the clinical decision-making process.

## CRedit authorship contribution statement

**Haizhou Wang, Fan Wang:** Conceptualization. **Yumei Ning, Kun Lin:** Data curation, Roles/Writing – original draft. **Yumei Ning, Haizhou Wang:** Formal analysis. **Jun Fang, Qiu Zhao, Fan Wang:** Funding acquisition. **Jun Fang, Qiu Zhao:** Project administration. **Yumei Ning, Yang Ding, Haizhou Wang:** Visualization. **Zhang Zhang, Xiaojia Chen, Fan Wang:** Writing – review & editing. **Yumei Ning, Kun Lin, Jun Fang:** Contributed equally to this work.

## Competing interests

The authors declare that they have no competing interests.

## Acknowledgements and funding

This work was supported by the Knowledge Innovation Program of Wuhan-Shuguang (2022020801020490), the Hubei Provincial Health and Family Planning Scientific Research Project (WJ2023M065), and the Program of Excellent Doctoral (Postdoctoral) of Zhongnan Hospital of Wuhan University (ZNYB2019003 and ZNYB2022003).

## Ethics approval and consent to participate

Not applicable.

## Consent for publication

Not applicable.

## Data Sharing Statement

Publicly available datasets were analyzed in this study. These data can be found here: <https://xenabrowser.net/datapages/>, <https://www.ncbi.nlm.nih.gov/geo/>, and <https://www.ebi.ac.uk/biostudies/arrayexpress>. All processed data and R codes used in this study can be obtained from the corresponding author on reasonable request.

## Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at [doi:10.1016/j.csbj.2023.04.016](https://doi.org/10.1016/j.csbj.2023.04.016).

## References

- [1] Vermeulen L, Snippert HJ. Stem cell dynamics in homeostasis and cancer of the intestine. *Nat Rev Cancer* 2014;14:468–80.
- [2] Rahib L, Smith BD, Aizenberg R, Rosenzweig AB, Fleshman JM, et al. Projecting cancer incidence and deaths to 2030: the unexpected burden of thyroid, liver, and pancreas cancers in the United States. *Cancer Res* 2014;74:2913–21.
- [3] Siegel R.L., Miller K.D., Jemal A. (2016) Cancer statistics, 2016. *CA: a Cancer Journal For Clinicians* 66.
- [4] Bijlsma MF, Sadanandam A, Tan P, Vermeulen L. Molecular subtypes in cancers of the gastrointestinal tract. *Nat Rev Gastroenterol Hepatol* 2017;14:333–42.
- [5] Adam RS, Blomberg I, Ten Hoorn S, Bijlsma MF, Vermeulen L. The recurring features of molecular subtypes in distinct gastrointestinal malignancies-A systematic review. *Crit Rev Oncol/Hematol* 2021;164:103428.
- [6] Salem ME, Puccini A, Xiu J, Raghavan D, Lenz H-J, et al. Comparative Molecular Analyses of Esophageal Squamous Cell Carcinoma, Esophageal Adenocarcinoma, and Gastric Adenocarcinoma. *Oncologist* 2018;23:1319–27.
- [7] Liu Y, Sethi NS, Hinoue T, Schneider BG, Cherniack AD, et al. Comparative Molecular Analysis of Gastrointestinal Adenocarcinomas. *Cancer Cell* 2018;33.
- [8] Anderson NM, Simon MC. The tumor microenvironment. *Curr Biol: CB* 2020;30:R921–5.
- [9] Pitt JM, Marabelle A, Eggermont A, Soria JC, Kroemer G, et al. Targeting the tumor microenvironment: removing obstruction to anticancer immune responses and immunotherapy. *Annals of Oncology: Official J Eur Soc Med Oncol* 2016;27:1482–92.
- [10] Duan Q, Zhang H, Zheng J, Zhang L. Turning cold into hot: firing up the tumor microenvironment. *Trends Cancer* 2020;6:605–18.
- [11] Deng B, Zhao Z, Kong W, Han C, Shen X, et al. Biological role of matrix stiffness in tumor growth and treatment. *J Transl Med* 2022;20:540.
- [12] Chen H, Cai Y, Chen Q, Li Z. Multiscale modeling of solid stress and tumor cell invasion in response to dynamic mechanical microenvironment. *Biomech Model Mechanobiol* 2020;19:577–90.
- [13] Hynes RO. The extracellular matrix: not just pretty fibrils. *Sci (N Y, N Y)* 2009;326:1216–9.
- [14] Levi N, Papisov N, Solomonov I, Sagi I, Krizhanovskiy V. The ECM path of senescence in aging: components and modifiers. *The FEBS J* 2020;287:2636–46.
- [15] Najafi M, Farhood B, Mortezaee K. Extracellular matrix (ECM) stiffness and degradation as cancer drivers. *J Cell Biochem* 2019;120:2782–90.
- [16] Ishihara S, Haga H. Matrix Stiffness Contributes to Cancer Progression by Regulating Transcription Factors. *Cancers* 2022;14.
- [17] Zhang W, Zhang S, Zhang W, Yue Y, Qian W, et al. Matrix stiffness and its influence on pancreatic diseases. *Biochim Et Biophys Acta Rev Cancer* 2021;1876:188583.
- [18] Liu C, Pei H, Tan F. Matrix stiffness and colorectal cancer. *Oncotargets Ther* 2020;13:2747–55.
- [19] Nersisyan S, Novosad V, Engibaryan N, Ushkaryov Y, Nikulin S, et al. ECM-receptor regulatory network and its prognostic role in colorectal cancer. *Front Genet* 2021;12:782699.

- [20] Wu G, Yang Y, Ye R, Yue H, Zhang H, et al. Development and validation of an ECM-related prognostic signature to predict the immune landscape of human hepatocellular carcinoma. *BMC Cancer* 2022;22:1036.
- [21] Liu C, Deng L, Lin J, Zhang J, Huang S, et al. Zinc Finger Protein CTCF Regulates Extracellular Matrix (ECM)-Related Gene Expression Associated With the Wnt Signaling Pathway in Gastric Cancer. *Front Oncol* 2020;10:625633.
- [22] Shen Y, Wang X, Lu J, Salfenmoser M, Wirsik NM, et al. Reduction of Liver Metastasis Stiffness Improves Response to Bevacizumab in Metastatic Colorectal Cancer. *Cancer Cell* 2020;37.
- [23] Bagaev A, Kotlov N, Nomie K, Svekolkin V, Gafurov A, et al. Conserved pan-cancer microenvironment subtypes predict response to immunotherapy. *Cancer Cell* 2021;39.
- [24] Moffitt RA, Marayati R, Flate EL, Volmar KE, Loeza SGH, et al. Virtual microdissection identifies distinct tumor- and stroma-specific subtypes of pancreatic ductal adenocarcinoma. *Nat Genet* 2015;47:1168–78.
- [25] Genomic Classification of Cutaneous Melanoma. *Cell* 2015;161:1681–96.
- [26] Chakravarthy A, Khan L, Bensler NP, Bose P, De Carvalho DD. TGF- $\beta$ -associated extracellular matrix genes link cancer-associated fibroblasts to immune evasion and immunotherapy failure. *Nat Commun* 2018;9:4692.
- [27] Li L, Wang X. Identification of gastric cancer subtypes based on pathway clustering. *NPJ Precision Oncology* 2021;5:46.
- [28] Aran D, Hu Z, Butte AJ. xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol* 2017;18:220.
- [29] Racle J, de Jonge K, Baumgaertner P, Speiser DE, Feller D. Simultaneous enumeration of cancer and immune cell types from bulk tumor gene expression data. *ELife* 2017;6.
- [30] Malta TM, Sokolov A, Gentles AJ, Burzykowski T, Poisson L, et al. Machine Learning Identifies Stemness Features Associated with Oncogenic Dedifferentiation. *Cell* 2018;173.
- [31] Hänzelmann S, Castelo R, Guinney J. GSEA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinforma* 2013;14:7.
- [32] Martínez-Jiménez F, Muiños F, Sentís I, Deu-Pons J, Reyes-Salazar I, et al. A compendium of mutational cancer driver genes. *Nat Rev Cancer* 2020;20:555–72.
- [33] Hafner M, Niepel M, Chung M, Sorger PK. Growth rate inhibition metrics correct for confounders in measuring sensitivity to cancer drugs. *Nat Methods* 2016;13:521–7.
- [34] Maeser D, Gruener RF, Huang RS. oncoPredict: an R package for predicting in vivo or cancer patient drug response and biomarkers from cell line screening data. *Brief Bioinforma* 2021;22.
- [35] Peña-Romero AC, Orenes-Piñero E. Dual Effect of Immune Cells within Tumour Microenvironment: Pro- and Anti-Tumour Effects and Their Triggers. *Cancers* 2022;14.
- [36] Zeltz C, Primac I, Erusappan P, Alam J, Noel A, et al. Cancer-associated fibroblasts in desmoplastic tumors: emerging role of integrins. *Semin Cancer Biol* 2020;62:166–81.
- [37] Palamaris K, Felekouras E, Sakellariou S. Epithelial to mesenchymal transition: key regulator of pancreatic ductal adenocarcinoma progression and chemoresistance. *Cancers* 2021;13.
- [38] Pepin KM, McGee KP, Arani A, Lake DS, Glaser KJ, et al. MR elastography analysis of glioma stiffness and -mutation status. *Ajnr Am J Neuroradiol* 2018;39:31–6.
- [39] Prasetyanti PR, Medema JP. Intra-tumor heterogeneity from a cancer stem cell perspective. *Mol Cancer* 2017;16:41.
- [40] Neeße A, Algül H, Tuveson DA, Gress TM. Stromal biology and therapy in pancreatic cancer: a changing paradigm. *Gut* 2015;64:1476–84.
- [41] Vennin C, Murphy KJ, Morton JP, Cox TR, Pajic M, et al. Reshaping the Tumor Stroma for Treatment of Pancreatic Cancer. *Gastroenterology* 2018;154:820–38.
- [42] Affo S, Yu L-X, Schwabe RF. The Role of Cancer-Associated Fibroblasts and Fibrosis in Liver Cancer. *Annu Rev Pathol* 2017;12:153–86.
- [43] Tauriello DVF, Palomo-Ponce S, Stork D, Berenguer-Llergo A, Badia-Ramentol J, et al. TGF $\beta$  drives immune evasion in genetically reconstituted colon cancer metastasis. *Nature* 2018;554:538–43.
- [44] Turley SJ, Cremasco V, Astarita JL. Immunological hallmarks of stromal cells in the tumour microenvironment. *Nat Rev Immunol* 2015;15:669–82.
- [45] García-Palmero I, Torres S, Bartolomé RA, Peláez-García A, Larriba MJ, et al. Twist1-induced activation of human fibroblasts promotes matrix stiffness by upregulating palladin and collagen  $\alpha 1(V)$ . *Oncogene* 2016;35:5224–36.
- [46] Zhang K, Grither WR, Van Hove S, Biswas H, Ponik SM, et al. Mechanical signals regulate and activate SNAIL1 protein to control the fibrogenic response of cancer-associated fibroblasts. *J Cell Sci* 2016;129:1989–2002.
- [47] Salmon H, Franciszewicz K, Damotte D, Dieu-Nosjean M-C, Validire P, et al. Matrix architecture defines the preferential localization and migration of T cells into the stroma of human lung tumors. *J Clin Invest* 2012;122:899–910.
- [48] Grasset EM, Bertero T, Bozec A, Friard J, Bourget I, et al. Matrix Stiffening and EGFR Cooperate to Promote the Collective Invasion of Cancer Cells. *Cancer Res* 2018;78:5229–42.
- [49] Gonzalez DM, Medici D. Signaling mechanisms of the epithelial-mesenchymal transition. *Sci Signal* 2014;7:re8.
- [50] Paauwe M., Schoonderwoerd M.J.A., Helderma Rfcp, Harryvan Tj, Groenewoud A, et al. (2018) Endoglin Expression on Cancer-Associated Fibroblasts Regulates Invasion and Stimulates Colorectal Cancer Metastasis. *Clinical Cancer Research: an Official Journal of the American Association For Cancer Research* 24:6331–44.
- [51] Tan Y, Tajik A, Chen J, Jia Q, Chowdhury F, et al. Matrix softness regulates plasticity of tumour-repopulating cells via H3K9 demethylation and Sox2 expression. *Nat Commun* 2014;5:4619.
- [52] You Y, Zheng Q, Dong Y, Xie X, Wang Y, et al. Matrix stiffness-mediated effects on stemness characteristics occurring in HCC cells. *Oncotarget* 2016;7:32221–31.
- [53] Tao B, Song Y, Wu Y, Yang X, Peng T, et al. Matrix stiffness promotes glioma cell stemness by activating BCL9L/Wnt/ $\beta$ -catenin signaling. *Aging* 2021;13:5284–96.
- [54] Lin C-H, Pelissier FA, Zhang H, Lakins J, Weaver VM, et al. Microenvironment rigidity modulates responses to the HER2 receptor tyrosine kinase inhibitor lapatinib via YAP and TAZ transcription factors. *Mol Biol Cell* 2015;26:3946–53.
- [55] Gao J, Rong Y, Huang Y, Shi P, Wang X, et al. Cirrhotic stiffness affects the migration of hepatocellular carcinoma cells and induces sorafenib resistance through YAP. *J Cell Physiol* 2019;234:2639–48.
- [56] Rice AJ, Cortes E, Lachowski D, Cheung BCH, Karim SA, et al. Matrix stiffness induces epithelial-mesenchymal transition and promotes chemoresistance in pancreatic cancer cells. *Oncogenesis* 2017;6:e352.
- [57] Hui L, Zhang J, Ding X, Guo X, Jiang X. Matrix stiffness regulates the proliferation, stemness and chemoresistance of laryngeal squamous cancer cells. *Int J Oncol* 2017;50:1439–47.
- [58] Dong B, Li S, Zhu S, Yi M, Luo S, et al. MiRNA-mediated EMT and CSCs in cancer chemoresistance. *Exp Hematol Oncol* 2021;10:12.
- [59] Colak S, Ten Dijke P. Targeting TGF- $\beta$  signaling in cancer. *Trends Cancer* 2017;3:56–71.
- [60] Sahai E, Atsaturov I, Cukierman E, DeNardo DG, Egeblad M, et al. A framework for advancing our understanding of cancer-associated fibroblasts. *Nat Rev Cancer* 2020;20:174–86.