

Senescence as a Dictator of Patient Outcomes and Therapeutic Efficacies in Human Gastric Cancer

Qiang Sun (✉ sunqiang1975@126.com)

Institute of Biotechnology, Beijing <https://orcid.org/0000-0002-5342-0637>

Lulin Zhou

Institute of Biotechnology

Zubiao Niu

Institute of Biotechnology

Yuqi Wang

Institute of Biotechnology

You Zheng

Second Military Medical University

Yichao Zhu

Institute of Biotechnology

Chenxi Wang

Institute of Biotechnology

Xiaoyan Gao

Institute of Biotechnology

Li-Hua Gao

Beijing Institute of Biotechnology

Wen Zhang

National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital

<https://orcid.org/0000-0003-3150-2751>

Kaitai Zhang

National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital

Gerry Melino

Torvergata Oncoscience Centre of Excellence University of Rome "Tor Vergata"

HONGYAN HUANG

<https://orcid.org/0000-0001-8340-6660>

Xiao-ning Wang

The Institute of Life Sciences <https://orcid.org/0000-0002-3762-7830>

Keywords: senescence, senescore, gastric cancer, prognosis, chemotherapy, immunotherapy

Posted Date: December 1st, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-855919/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Cell Death Discovery on January 10th, 2022. See the published version at <https://doi.org/10.1038/s41420-021-00769-6>.

Abstract

Senescence is believed to be a pivotal player in the onset and progression of tumors as well as cancer therapy. However, the guiding roles of senescence in clinical outcomes and therapy selection for patients with cancer remains obscure, largely due to the absence of a feasible senescence signature. Here, by integrative analysis of single cell and bulk transcriptome data from multiple datasets of gastric cancer patients, we uncovered senescence as a veiled tumour feature characterized by senescence gene signature enriched, unexpectedly, in the non-cancerous cells, and further identified two distinct senescence-associated subtypes based on the unsupervised clustering. Patients with the senescence subtype had higher tumor mutation loads and better prognosis as compared with the aggressive subtype. By the machine learning, we constructed a scoring system termed as senescore based on 6 signature genes: ADH1B, IL1A, SERPINE1, SPARC, EZH2 and TNFAIP2. Higher senescore demonstrated robustly predictive capability for longer overall and recurrence-free survival in 2290 gastric cancer samples, which was independently validated by the multiplex staining analysis of gastric cancer samples on the tissue microarray. Remarkably, the senescore signature served as a reliable predictor of chemo- and immunotherapeutic efficacies, with high-senescore patients benefited from immunotherapy while low-senescore patients were responsive to chemotherapy. Collectively, we report senescence as a heretofore unrecognized hallmark of gastric cancer that impacts patient outcomes and therapeutic efficacy.

Introduction

Gastric cancer (GC) is the third leading cause of cancer death and the fifth most common cancer globally¹. The chemotherapy is one of the major therapeutic strategies for GC patients. However, there is considerable heterogeneity in the treatment response and clinical outcomes among cancer patients despite in similar clinical or pathologic conditions², which underscores the need for novel predictive markers other than clinical stages and pathohistological classifications. With the rapid advance in the high-throughput technology, molecular subtypes of GC were established by The Cancer Genome Atlas (TCGA) including microsatellite instable (MSI), genomically stable (GS), chromosomal instability (CIN), Epstein-Barr virus (EBV) associated, and hypermutated-single-nucleotide variants (HM-SNV)³, as well as by the Asian Cancer Research Group (ACRG) comprising microsatellite stable (MSS)/epithelial-mesenchymal transition (EMT), MSI, MSS/p53+, and MSS/p53-⁴. However, these defined classifications face huge obstacles in the clinical translation and limitations in relevant cancer patients.

Cancer immunotherapy like immune checkpoint inhibitors (ICIs) elicits durable antitumor responses in multiple solid cancers. However, only a fraction of patients with cancer respond to immunotherapy, calling for biomarkers for better clinical management⁵. Although it was revealed that a wide spectrum of tumour features, such as genomic instability described as MSI or tumour mutational burden (TMB) status and tumour-infiltrating immune cells in tumour microenvironment (TME), were associated with immunotherapy sensitivity in preclinical and clinical settings⁵⁻⁷, it is still a daunting challenge to predict

the treatment responses of chemotherapy and immunotherapy simultaneously, and to make the appropriate therapeutic options.

Cellular senescence, defined as the state of cell-cycle arrest, can be triggered by severe interior or exterior insults such as oncogenic activation or chemotherapeutic DNA damage^{8,9}. Cellular senescence was believed to be a failsafe program for the organism by excluding deleterious cells from further expansion, and thereby suppresses tumorigenesis in a cell-autonomous setting¹⁰⁻¹². On the other hand, senescent cells can secrete a plethora of cytokines and growth factors including interleukin-6 (IL-6) and IL-8, termed as senescence-associated secretory phenotype (SASP), whereby establish an immunosuppressive, inflammatory and catabolic microenvironment to promote tumor growth and resistance to chemotherapy¹²⁻¹⁵. Conversely, the SASP factors can also trigger senescence in an autocrine or paracrine fashion to impose an anti-tumor effect^{16,17}. Therefore, it is imperative to elaborate the impacts of the senescence on tumors for better cancer prevention and therapy. Furthermore, there exist a wide spectrum of attributes that distinguish senescent cells from their normal counterparts, such as prominent changes in the cell size and morphology, increased expression of the CDK inhibitors CDKN2A and enhanced DNA damage response (DDR), but the unique and specific markers for senescent cells are still absent^{8,18}. In fact, though preclinical studies involving the cell experiments and animal models have revealed the extensive effects of senescence on cancer, our knowledge about the senescent characteristic in tumor samples of patients is still very rudimentary. In particular, whether the senescent feature in patients with cancer could serve as a biomarker to guide clinical prognosis and treatment remain unknown.

In this study, we performed the comprehensive analysis of senescent features within TME in multiple GC cohorts, and further identified the senescence subtypes and dissected its association with clinical and molecular features, as well as the immune cell infiltration and signaling pathway activation. Finally, we demonstrated that senescore, a senescence scoring system of tumor, exhibited robust predictive powers not only for the patients outcomes, but also for the efficacies of different therapeutic strategies including adjuvant chemotherapy and immunotherapy (Fig. 1).

Results

Identification of senescence as a hallmark of gastric cancer

To elucidate the senescence character of GC, we comprehensively explored the transcriptomic data in multiple datasets of GC. In TCGA-STAD cohort, the results of gene set enrichment analysis (GSEA) showed that those biological pathways indicative of cellular senescence, including cellular senescence (NES=1.63, $p < 0.01$), DNA damage/telomere stress induced senescence (NES=1.98, $p < 0.01$), and oxidative stress induced senescence (NES=1.58, $p < 0.01$), were significantly enriched in the GC over normal gastric tissues (Fig. 2A). And Gene Ontology (GO) about cellular senescence was also significantly enriched in GC compared with normal tissues (Fig. S1A). Similarly, consistent results for senescence were observed in the another independent cohort (Fig. 2B, Fig. S1B). Moreover, the gene set variation analysis between

GC and their paired adjacent normal tissues in 3 datasets showed GC qualified with markedly activation in multifarious senescence pathways and GO terms (Fig. 2C). These results suggested senescence as a prominent hallmark of GC.

To investigate the genes impacting senescence in GC, we firstly screened the differently expressed genes (DEGs) between GC and no-tumor adjacent tissues in TCGA-STAD dataset. It was shown that 3225 DEGs were identified, including 1455 up-regulated and 1770 down-regulated genes (Fig. S1C, D), and the up-regulated genes were markedly enriched in the cellular senescence pathway (Fig. S1E). Then, we acquired senescence genes based on all the gene sets associated with senescence in Molecular Signatures Database¹⁹ (Table S1). Combined the DEGs with senescence genes, 103 senescence-associated genes in GC were identified. GO and pathway enrichment analysis of these genes was conducted, showing significantly enriched biological processes comprising aging, cell cycle and cellular senescence (Fig. 2D, E). As shown in Fig. 2F, the interaction network between top five enriched pathways and genes encompassed the specific marker for cellular senescence (CDKN2A) and immune-associated genes (e.g. CXCL8 and IL6).

Single cell profiling of senescence within tumour microenvironment in gastric cancer

To further dissect the senescence characteristics within GC tissues at single cell level, we investigated the single cell transcriptome of 19042 cells from normal and cancerous gastric tissues (Fig. 3A, B, Fig. S2A). In agreement with the results from the bulk analysis above, more cells with significant activation of cellular senescence pathway were identified in GC as compared with those in normal gastric tissues (Fig. 3C, Fig. S2B). Interestingly, in GC, certain cell types consisting of endothelial cells, enteroendocrine cells, proliferative cells and macrophages as well as fibroblasts, but not the cancer cells, are highly enriched in various senescence pathways (Fig. 3D), suggesting these cell types, rather than the cancer cells, may primarily contribute to the senescence features in GC.

To identify the key genes associated with senescence in GC, we integrated significantly different expression genes (DEGs), the survival-associated genes and senescence genes derived from senescence gene sets, and finally established the core senescence genes (CSGs) of GC (Fig. 3E, F). At single cell level, CSGs were expressed with higher proportions in cells of TME such as endothelial cells, fibroblasts and macrophages (Fig. 3G), which is in line with above pathway analysis.

Furthermore, we assessed the incidence of copy number variations (CNV) and somatic mutations of 17 CSGs in GC using genomic data. It was indicated that *COL1A2* gene exhibited the highest mutation frequency followed by *ABCA8*, while the investigation of CNV alteration frequency showed that a prevalent CNV amplification occurred in *COL1A2*, *C7*, *ABCA8*, as well as *SERPINE1* gene (Fig. 3H and Fig. S2C, D). Although the gene-expression changes were associated with the genomic transformation, especially the alteration of CNV (Fig. S2C,E), the complexity of genomics and heterogeneity of tumor led to the incomplete consistency in the landscape of genomic and expressional alterations, highlighting that expression characteristics of CSGs played a considerably pivotal role in GC.

The identification of senescence subtypes based on the core senescence genes

In order to decipher the underlying biological characteristics of the senescence phenotype, the unsupervised hierarchical clustering analysis was performed firstly in the discovery cohort of TCGA-STAD based on the expression of CSGs. Two senescence-associated clusters were identified (Fig. 4A), and principal component analysis (PCA) indicated the significant distinction existed on the transcriptional profile between the two subtypes (Fig. 4B, Fig. S3A). Except for *EZH2*, *TNFAIP2* and *IL1A*, the remaining genes exhibited higher expression in cluster 1 than those in cluster 2 (Fig. S3B). Furthermore, functional enrichment analysis showed that cluster 2 displays significantly higher activation in senescence pathways, suggesting cluster 2 was associated with senescence subtype (Fig. 4C). Survival analysis showed that cluster 2 was associated with longer overall survival (OS) (Fig. 4D), which was independently validated in a pooled cohort of 1484 patients (Fig. S3C,D). Interestingly, the higher mutation density was observed in cluster 2 in comparison with cluster 1 in TCGA data set (Fig. 4E).

Moreover, the senescence subtypes were parsed in the clinical and molecular features in TCGA cohort. The results showed that majority of patients with early clinical stage (stage I) were in cluster 2 (Fig. S4A). And no significant distribution difference was found between groups stratified by the age of patients (Fig. S4B), which was probably resulted from the little age gap among patients (Table S2). We also found that tumors in cluster 1 presented poorer differentiation and were enriched in the diffuse and signet-ring histological classification (Fig. S4C), suggesting cluster 1 was associated with the more aggressive subtype. With regards to the molecular subtypes of TCGA project³, the genomically stable (GS) subtype was mainly concentrated in the cluster 1, while the subtypes of microsatellite instability (MSI) and Epstein-Barr virus (EBV)-positive substantially appeared in the cluster 2 (Fig. S4D). In consistent with the molecular subtype, cluster 2 was higher in the proportion of MSI tumors by the MSI status analysis and the clonal deletion score for quantificational chromosomal instability (CIN) than those in cluster 1 (Fig. S4E,F).

Additionally, we analyzed the canonical biological processes associated with TME, indicating that the cluster 2 displayed activation of DNA damage response (DDR), antigen processing and presentation (APM), which may result from the higher mutation load (Fig. 4F)²⁰. Furthermore, we then explored the characteristics of the immune cell infiltration in 1484 GC patients in the validation cohort 1. The results showed that cluster 1 showed significant increases in the infiltration of anti-tumor immune cells such as the activated CD8⁺ T cells and central memory CD4⁺ T cells, as well as pro-tumor immune cells like the regulatory T cells (Treg), and myeloid-derived suppressor cells (MDSC). By contrast, cluster 2 appeared high infiltration of activated CD4⁺ T cells, CD56 bright natural killer cells (NK), and memory B cells (Fig. 4G). These results indicated that the senescence subtypes were associated with distinct features of TME.

Construction of a senescence scoring model

The senescence phenotype acted non-negligible roles in shaping the TME and markedly affected the clinical prognosis. However, these analyses were only based on the patient population and could not accurately predict the senescence character in individual patients. Therefore, we sought to construct a scoring system to quantify the senescence level in tumour tissues for individual patient with GC. In the TCGA-STAD cohort, we constructed a senescence scoring model, termed “senescore”, by using 6 senescence signature genes (ADH1B, IL1A, SERPINE1, SPARC, EZH2 and TNFAIP2) identified with the LASSO Cox regression algorithm (Fig. S5A, S5B). The survival analysis demonstrated patients with high senescore were correlated with better outcomes in TCGA-STAD cohort (log-rank test, $P < 0.0001$; median OS 19.6 months versus 56.2 months, Fig. 5A and Table S3). Furthermore, when the senescore was evaluated as a continuous variable with the multivariate Cox regression model, the senescore was determined to be an independent and robust prognostic factor (HR, 0.14; 95%CI, 0.065–0.32; $P < 0.001$; Fig. 5B). More importantly, when patients were stratified with TNM clinical stages, the senescore successfully separated patients with distinct clinical prognosis at the same disease stage (Fig. 5C). These findings suggested that the senescore reserved its prognostic relevance even after classic clinicopathologic prognostic features have been taken into accounts. Therefore, for better clinical utility, we integrated senescore with the age and clinical stages, two independent prognostic factors identified above, to generate a comprehensive nomogram (Fig. 5D). The prognostic capacity of the nomogram was also demonstrated by the area under the curve (AUC) of the time-dependent receiver operating characteristic (ROC) curve with a value of 0.794 at 5 years (Fig. S5C). Furthermore, the calibration curves of the model for the possibility of OS at 3 years and 5 years closely approximated the observed estimates, manifesting the accurate predictive ability (Fig. S5D, S5E).

To better illustrate the feature of senescore, the overview of the connection among the senescence subtypes, molecular subtypes, immune subtypes and the senescore of individual patients in TCGA cohort was depicted by the alluvial diagram (Fig. 5E). And the correlations between the senescore and varying biological processes suggested that senescore was positively associated with the activation of DDR, senescence, and DNA damage repair pathway (Fig. 5F). The senescore was significantly higher in cluster 2 subtype than that in cluster 1 (Fig. 5G). Whereas for the established molecular subtypes, the GS subtype had the lowest median senescore, while the EBV subtype and the MSI subtype showed significantly correlated with higher senescore (Fig. 5H). Similarly, the senescore was also associated with the MSI status, corresponding microsatellite instability-high (MSI-H) to the highest senescore (Fig. S5F). Furthermore, the senescore was positively correlated with the mutation density (Fig. 5I). Additionally, senescore signature was associated with histological classification, showing patients with intestinal GC had highest senescore (Fig. S5G). These findings suggested that the senescore was closely associated with clinical characteristics of GC patients and may serve as a prognostic biomarker.

Validation of senescore to effectively predict patient outcomes

To confirm the robustness of the senescore in different populations, we further explored ACRG cohort that was provided with the comprehensive clinical information. Similar with the TCGA cohort, the survival

analysis revealed that the high senescore was associated with the longer OS as well as the RFS ($p < 0.001$; hazard ratio, 0.12 [95%CI, 0.04-0.33]) (Fig. 6A). And the prognostic nomogram demonstrated an impressive predicting ability for OS (1-year AUC = 0.81, 3-year AUC = 0.77, 5-year AUC = 0.75) as indicated by the time-dependent ROC curve analysis (Fig. 6B). Moreover, the significant differences of the senescore were also observed among molecular subtypes in ACRG cohort. Rather, the MSI subtype showed the highest senescore, while the MSS/EMT subtype had the lowest senescore (Fig. 6C). In addition, the histological classification also displayed the obvious differences of the senescore values (Fig. S6A), which was consistent with the results in TCGA-STAD cohort. Additional validation of the senescore model was further confirmed in the validation cohort1 (n=1484) for OS and in the validation cohort2 (n=999) for RFS, as well as all GC patients in our study (Fig. 6D,E, Fig. S6B). Next, the prognostic power of senescore was examined in a wide spectrum of gastrointestinal tumours, and significantly validated in colorectal cancer and oesophageal cancer, respectively, suggesting senescore signature may be conserved in other types of cancer (Fig. S6C-G).

Senescore predicts the efficacies of chemotherapy and immunotherapy

Considering that cellular senescence could be induced by DNA damage drug, we explored whether senescore, the senescence scoring model, could predict the treatment efficiency of chemotherapy. In ACRG cohorts with available adjuvant chemotherapy information, patients were divided into high and low senescore group and the difference in OS and RFS was independently assessed. Adjuvant chemotherapy was found to markedly improve the overall survival rate and recurrence-free survival rate in patients with low senescore ($P = 0.00037$), while patients with the high senescore showed only a moderate benefit from adjuvant chemotherapy ($P=0.08$, Fig. 6F, S7A). Furthermore, patients with GC in another three cohorts of adjuvant chemotherapy were categorized into high- and low-senescore group. And significant survival difference was observed in the low-senescore group, but not in the high-senescore group (Fig. 6G, S7B), which was consistent with the results in ACRG cohort. These results suggest that the patients with low-senescore were more inclined to gain benefits from the chemotherapy.

However, in contrast to chemotherapy, patients with high senescore, especially the highest quartile senescore, exhibited significantly prolonged OS in an independent anti-PD-L1 therapy cohort (n=348) (Fig. S7C, S7D)²¹. In addition, patients responsive to anti-PD-L1 therapy displayed higher senescore (Fig. S7E) and senescore in this immunotherapy cohort was positively related with tumor mutation load (Fig. S7F), suggesting patients with high senescore was associated with better immunotherapy efficiency. Similarly, patients with high senescore in a small anti-PD1/PD-L1 cohort (n = 27)²² displayed longer progress free survival (PFS) (Fig. 6H), as well as higher overall response (50% vs 21%) compared with low senescore group (Fig. 6I). Similar findings were independently confirmed in a cohort of large size (n = 399), in which the patients of urothelial tumors²¹ and melanoma²³ receiving anti-PD1/PD-L1 therapy were combined, where the high-senescore group displayed a significantly longer OS (Fig. 6J), and a higher proportion of overall response (31% vs 19%) (Fig. 6K). These results indicated that senescore might serve as a potential biomarker helping the clinicians to select appropriate patients for either chemotherapy or immunotherapy.

Validation of senescore at translational level

To further validate our model at translational level for the purpose of clinical pathology, a sophisticated method of multiplex immunofluorescence histochemistry, which allows simultaneous detection of multiple target proteins, was employed to analyse the protein expression of the 6 senescence signature genes identified above on a GC tissue microarray (TMA, Fig. 7A). Among these senescence signatures, four molecules, each by their own, were associated with clinical prognosis (Fig. S8A-F). Consistent with transcriptional analysis above, higher senescore in TMA cohort was significantly associated with longer OS (HR = 0.28, 95%CI: 0.17-0.45; $p < 0.0001$) (Fig. 7B). Furthermore, the ROC curve was conducted to evaluate the sensitivity and specificity of the senescore, and it illustrated that the AUC values was 0.882, 0.854 and 0.878 for 1, 3 and 5 years, respectively (Fig. 7C). In addition, ROC curve analysis also demonstrated the forceful predictive ability of the prognostic nomogram integrating senescore with clinical stage and age (Fig. 7D). More importantly, the senescore displayed stronger predictive value for OS in comparison with TNM stages, while the nomogram model combining the clinical stage and age with senescore only slightly elevated the predictive accuracy of the patients' prognosis as compared with senescore alone (Fig. 7E). Collectively, senescore is promising in predicting clinical prognosis by feat of the immunofluorescence histochemistry strategy.

Discussion

In this study, we identified senescence as a hallmark of GC that is associated with distinctive TME and genomic alterations. Therefore, the identification of the senescence subtypes provides an insightful perspective on tumors from another angle. For example, The senescence subtype(cluster 2) showed the features of MSI as well as CIN. It is explained that tumors with genomic instability have more somatic mutations, which can induce cellular senescence by the pathway of DNA damage response²⁴. Finally, it was revealed that cluster 2 presented significant survival advantage, highlighting the activation of cellular senescence pathway may play a beneficial role in the clinical prognosis of patients with GC, and likely other types of cancers as well²⁵.

Furthermore, we established the senescence scoring algorithm to evaluate the senescence level of tumor tissues, termed as senescore. Our results demonstrated that the senescore is of importance in clinical relevance. Integrated analysis revealed that the senescore is a robust prognostic biomarker in a vast number of cancer patients. More importantly, the senescore is able to serve as an independent prognostic factor. Furthermore, the senescore is associated with different established molecular subtypes, histological subtypes, and genomic alterations. For example, high senescore is related with the MSI-H status, EBV infection and higher TMB. Several studies have demonstrated that patients with MSI-H and EBV-positive tumors as well as high TMB are more inclined to dramatic responses to immunotherapy such as immune checkpoint inhibitors²⁶⁻²⁸. Consistent with our analysis above, it was indicated that high senescore is more likely to gain benefit from the immunotherapy. Moreover, the technology of multiplex immunofluorescence histochemistry demonstrated a good validation of senescore in robustly

predicting patient outcomes, which is promising in bridging the senescence scoring system to the clinical practice in near future.

More importantly, subset analysis of patients with available chemotherapy data strongly suggested that GC patients with low senescore was more cline to gain benefit from the chemotherapy. Previous studies have demonstrated that cellular senescence could increase the drug resistance and side effects of the chemotherapy^{29–31}, suggesting high senescore associated with the activation of the senescence program that might blunt the desirable effects of the chemotherapy.

In conclusion, this is the first comprehensive senescent analysis of GC, leading to the identification of senescence subtypes that are associated with significantly different tumor microenvironment and survival outcomes. Our analysis demonstrated that the senescore can be used to identify not only patients at different risk levels for the OS as well as RFS, but also patients who would benefit from adjuvant chemotherapy or immunotherapy therapy. Impressively, the association between the senescore and immunotherapy effects, though being relatively weak due to the limit of available patient samples, was significant, which encourages further evaluation in a larger size of patients taking immunotherapy. Nevertheless, the validation of our findings in a wide spectrum of patient cohorts, and the findings that the senescence features reflect biological and clinical characteristics associated with sensitivity or resistance to the therapy, would pave a way for developing more rational therapy recommendations and promoting personalized cancer therapy.

Materials And Methods

Data collection and processing

We retrospectively acquired the sample information of 2290 patients of gastric cancer (GC) from 12 eligible GC datasets from The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) database. In addition, transcriptomic and clinical data of 1324 patients with other digestive system tumors including colorectal cancer, esophageal cancer, liver cancer and pancreatic cancer were obtained from TCGA database. And an independent dataset of 579 patients with colon cancer was download from GEO database. Overall, the information of all the datasets enrolled in this study was summarized in Table S4.

For TCGA dataset, the raw count expression data and the normalized RNA sequencing data (FPKM value) were downloaded from TCGA Data Portal using the TCGAbiolinks R package³². Then, FPKM values of the normalized RNA sequencing data were transformed into transcripts per kilobase million (TPM) values for analysis. Moreover, the microarray gene-expression data from GEO database were obtained and processed through two strategies. For microarray data of Affymetrix platform, we downloaded the raw “CEL” files and adopted the Robust Multiarray Average (RMA) algorithm for background adjustment and quantile normalization, and the median polish algorithm for final summarization of oligonucleotides per transcript in the affy R package. For microarray data from other platforms, the processed normalized

matrix data were directly downloaded by the GEOquery R package³³. Batch effects from technological biases were corrected using the ComBat algorithm of sva package when necessary³⁴. All the R packages in this study were performed in the R software (version 3.6.1, <https://www.r-project.org>).

Patient population and clinical information

In this study, we retrospectively collected 2290 patients of GC from the curated public databases for survival analysis. The TCGA-STAD cohort as the exploratory cohort included 374 patients with detailed clinical information (TableS2 and S3). The validation cohorts embodied an independent ACRG cohort (GSE66229) and two large pooled cohorts (validation cohort 1 and validation cohort 2). The validation cohort1 (n=1484) used as overall survival (OS) analysis was integrated with identifiers GSE26899, GSE26901, GSE28541, GSE13861, GSE66229, GSE15459, GSE34942, GSE84437, GSE57303 and GSE29272. And the validation cohort2 (n=999) as recurrence-free survival (RFS) analysis was comprised of identifiers GSE26253, GSE26899, GSE26901, GSE13861 and GSE66229. OS was established as the time from surgery to death or to the last follow-up time, and RFS was defined as the date from surgery to the first verified recurrence. Clinical data of patients in these cohorts was summarized in Table S2.

Furthermore, adjuvant chemotherapy data were available for the patients in the ACRG cohort including 80 patients receiving postoperative adjuvant chemotherapy⁴. With regard to another validation cohort of the adjuvant chemotherapy, 267 patients who underwent gastrectomy as primary treatment at 3 cohorts (KUGH cohort under GSE26899, KUCM cohort under GSE26901 and YUSH cohort under GSE13861 datasets) were enrolled, including 155 patients who had received standard adjuvant chemotherapy (either single-agent 5-fluorouracil or a combination of 5-fluorouracil and cisplatin/oxaliplatin, doxorubicin, or paclitaxel) (Table S5)³⁵. We collected 3 independent datasets of patients treated with immunotherapies with available transcriptomics data (Table S6). For the Mariathasan, et al. anti-PD-L1 therapy cohort, the data of 348 patients with metastatic urothelial tumors treated with the anti-PD-L1 agent were obtained from IMvigor210CoreBiologies R package (<http://research-pub.gene.com/IMvigor210CoreBiologies/>)²¹. The raw count data was normalized by implementing trimmed mean of M-values (TMM) method and transformed with the “voom” function in the limma package³⁶. For another anti-PD1/PDL1 therapy cohort of lung cancer (n=27), transcriptomic data was downloaded from GEO database under GSE135222 and only the clinical information of progress-free survival (PFS) was available²². Additionally, in nivolumab treatment cohort, pre/on therapy biopsies of 68 patients were analyzed by transcriptome sequencing, which was available on GEO database (GSE91061). We selected only the pre-treatment samples with clinical information for further analysis²³. Tumor response for all patients was defined by RECIST v1.1.

The corresponding clinical data from these patients were retrieved and manually arranged through the following two main methods:(1) utilizing the GEOquery R package or directly downloading on the GEO website through the corresponding accession number; (2) from the supplementary materials of the corresponding literature. The related authors were contacted for further information when necessary. Details of ethical approval and informed consent for all studies can be found in corresponding publications.

Identification of differentially expressed genes and functional annotation

The differentially expressed genes (DEGs) between GC and normal gastric tissues were screened via the edgeR R package³⁷. The significant criteria for determining DEGs was set as the false discovery rate (FDR) < 0.05 and the |fold change|>2. Gene annotation enrichment analysis for Gene ontology (GO) and Reactome pathway using the clusterProfiler R package³⁸ was conducted on senescence genes. Gene Set Enrichment Analysis (GSEA) was performed through clusterProfiler R package. And the genes were ordered according to the fold change of gene expression between the GC and normal tissue.

Genomic data analysis

The somatic mutation data was acquired from TCGA mutations R package³⁹. Besides, copy number variations (CNV) and somatic mutation data in the meta-cohort was acquired from the cBio Cancer Genomics Portal (cBioPortal, <http://www.cbioportal.org/>)^{40,41}. The mutation landscape was depicted using the functions of maftools R package⁴². Genomic features of TCGA datasets, including clonal detection score, aneuploidy score, mutation density et al, were obtained from the published research³. And data about the molecular subtypes of TCGA cohorts were also retrieved where available⁴³ (Table S3).

Single cell RNA-seq data analysis

In order to further portray senescence microenvironment in GC at single cell level, we utilized the single cell RNA (scRNA) data of gastric tissues previously published including non-atrophic gastritis (NAG), chronic atrophic gastritis (CAG), intestinal metaplasia (IM), and GC⁴⁴. The analytical pipeline of scRNA data as previously described⁴⁵. Briefly, the raw count matrices of scRNA data was processed by the Seurat R package in R software⁴⁶. Then followed by quality control to filter low-quality cells, cells expressed lower than 400 genes or higher than 7000 detected genes, and more than 20% of all gene counts mapped to mitochondrial or ribosomal genes were removed as mentioned by the corresponding authors⁴⁴. Then, the single-cell gene expression data was normalized by using “NormalizeData” function with setting normalization method “LogNormalize” in the R package Seurat. We then utilized “ScaleData” function to regress out cell-cell variation driven by batch, the gene counts, as well as mitochondrial and ribosomal gene expression.

Furthermore, the highly variable genes were selected for the dimension reduction by “FindVariableFeatures” function in the Seurat package. Then, we used the “RunPCA” function in Seurat to perform the principle component analysis (PCA). To establish the numbers of principle components representing the variance of cells, the “ElbowPlot” function was implemented, and 20 significant principal components were identified for further analysis.

We then used the “FindNeighbors” function in Seurat to assemble cells based on a graph-based clustering approach. Then, the “FindCluster” function with the resolution parameter 2 was implemented to identify the number of cell clusters. We utilized the “FindAllMarkers” function to find differentially expressed genes by comparing each cluster of cells with all other cells, and annotated cell clusters based on the expression of curated known cell markers.

Unsupervised clustering for the senescence subtypes

To characterize senescence signature, Unsupervised clustering analysis was applied to identify distinct clusters based on the expression of CSGs. We adopted the function of the pheatmap R package based on the Euclidean distance and Ward's linkage method to classify patients unsupervisedly. The number of clusters and their stability were determined by the NbClust package, which provided 30 indices for determining the number of clusters and proposed the best clustering scheme⁴⁷.

Characterization of tumor microenvironment

To quantify the relative abundance of each cell infiltration in the GC, we used the single-sample Gene Set Enrichment Analysis (ssGSEA) algorithm. The gene sets for analyzing each immune cell type were obtained from the study of Charoentong⁴⁸. The enrichment scores calculated by ssGSEA analysis were utilized to represent the relative abundance of each type of immune cell^{49, 50}.

To investigate the difference of the biological process in different samples, Gene set variation analysis (GSVA) analysis was performed using GSVA R packages. GSVA is a gene set enrichment method to estimate the variation in the pathway and biological process activity among different samples in a non-parametric and unsupervised manner⁵¹. The gene sets about senescence and other biological processes were downloaded from Molecular Signatures Database (MSigDB) of the Broad Institute for running GSVA analysis (Table S1)⁵². Additionally, the gene sets that contained genes associated with tumor microenvironment, including the angiogenesis signature⁵³, antigen processing machinery (APM), DNA damage repair response (DDR), epithelial mesenchymal transition (EMT), pan-fibroblast TGF β response signature (Pan F-TBRS) and TGF β pathway, were retrieved from a previous research²¹.

Development and validation of the senescence scoring and the prognostic nomogram

To dissect the senescence signature of individual tumor for better clinical utility, we developed a senescence score termed as senescore through least absolute shrinkage and selection operator (LASSO) machine learning algorithms. The specific process was as follows. The core senescence genes (CSGs) were put into the Cox regression model with LASSO penalty for analysis using the glmnet R package⁵⁴. Genes of the model were established by the appropriate value of the penalty parameter λ , which were determined by ten-fold cross-validations in TCGA-STAD cohort. The senescence risk score model was produced by integrating the expression level of selected genes and their corresponding coefficients derived from the prognostic model analyses⁵⁵, as follows: senescence risk score = (0.0394* ADH1B) +

(0.0515* IL1A) + (-0.1386*TNFAIP2) +(-0.0473*EZH2) +(0.0100* SPARC) +(0.1513* SERPINE1). In order to make a positive consistency of senescence score and senescence features, we reformulated the senescence risk score as the senescore⁵⁶:

$$\text{Senescore} = e^{-\sum_{i=1}^n \beta_i * x_i}$$

β_i represents the regression coefficient of each gene in Lasso regression model, X_i shows the gene expression value of each gene.

The prognostic nomogram was generated based on the senescore using the rms R package and externally validated in the GSE66229 cohort. We conducted 3-, 5-year OS calibrations to determine the predictive accuracy of the nomogram model. The concordance index (C-index) and receiver operating characteristic (ROC) curve were also used to evaluate the predictive accuracy of the model, and the time dependent ROC curve was made using survivalROC R package.

Multiplex immunofluorescence immunohistochemistry and imaging

A tissue microarray spotted with tumor samples from 90 GC patients (HStmAde180Sur05) was purchased from Shanghai Outdo Biotech Co. Ltd. A total of 190 cores on the slide consisted of 90 cases of breast cancer tissue and paired normal breast tissue. The detailed clinical information was summarized in table S7. All tissues were collected in accordance with the ethical standards with the donor being informed completely and with their consent, from National Human Genetic Resources Sharing Service Platform: 2005DKA21300.

For multiplex immunofluorescence staining⁵⁷, the Opal 7-color manual IHC kit (PerkinElmer, cat. No. NEL811001KT) was used according to the manufacturer's instructions, and the molecules panel, which consisted of six antibodies including anti-IL1A (Proteintech, cat.No.16765-1-AP, 1:250 dilution), anti-ADH1B (Proteintech, cat.No.17165-1-AP, 1:500 dilution), anti-SERPINE1 (ABclonal, cat. No. A14758, 1:200 dilution), anti-SPARC (ABclonal, cat. No. A14494, 1:200 dilution), anti- EZH2 (ABclonal, cat. No. A19577, 1:200 dilution) and anti-TNFAIP2 (Signal Antibody, cat.No.40163, 1:200 dilution), was conducted on the same slide. In brief, the slides were baked at 65°C for 1 hour. Deparaffinization with xylene for 15 minutes three times was followed by rehydration in an ethanol series. Antigen retrieval was performed in the EDTA antigen retrieval buffer (pH 8.0) using microwave treatment, and the nonspecific binding was blocked with the blocking buffer for 10 min at room temperature. After that, the slide was stained with antigen-specific primary antibodies for 1 hour at room temperature or overnight at 4°C, following by Opal Polymer Horseradish Peroxidase (HRP) for 10 min at room temperature. Then a specific fluorochrome from the Opal 7-color manual IHC kit for each primary antibody was used to visualize the respective antigens. For each consecutive antibody staining, the antigen retrieval step, based on microwave treatment like the prior operation, was performed, which allowed for the removal of prior primary and secondary antibodies while the fluorophore remained covalently bonded to the tissue antigen.

The stained slides were scanned using the Vectra Polaris Automated Quantitative Pathology Imaging System (PerkinElmer) by a 20×objective lens. Inform automated image analysis software (PerkinElmer) was used for batch analysis of multispectral images based on the specific algorithms as previously described⁵⁸. Multispectral images were decomposed into their single component by spectral unmixing using a digital spectral library consisting of spectral profiles of each of the fluorophores.

Due to the stripping the slides and limitation of scanning technique, 80 tissue samples were involved in final analysis. The IHC images of each molecular staining results were scored twice by the clinician. The protein scoring system was based on the intensity and extent of the staining as previously described⁵⁹. The staining intensity was classified as 0 (negative), 1 (weak), 2 (moderate) or 3 (strong). The staining extent was dependent on the percentage of positive area: 0 (< 5%), 1 (5%-25%), 2 (26%-50%), 3 (51%-75%) or 4 (>75%). Each protein score was based on the stain intensity and extent scores, and then normalized with a z-score. Finally, the senescore was calculated on the basis of the protein expression scores of IL1A, ADH1B, TNFAIP2, SERPINE1, SPARC and EZH2 (Table S7)

Statistical analysis

The distribution of the variables was tested by the Shapiro-Wilk test. For the comparison of two groups, statistical significance of variables was analyzed by the unpaired Student t test or the Wilcoxon rank-sum test. The Kruskal-Wallis test was applied to evaluate significant differences when comparing more than two groups. Correlations coefficients between CSGs were computed by Spearman at the transcriptomic level.

The prognosis genes were determined by univariate Cox regression analysis. The survival curves for the prognostic analysis were generated via the Kaplan-Meier method and the log-rank test was used to compare the statistical differences in survival between two groups. To analyze the significance of senescore to predict clinical prognosis, the cutoff values of senescore were establish by the median of the data in each separate data set. Alternatively, to attenuate the effect of the group method on prognosis, the cut-off point was also determined using “surv-cutpoint” function in the survminer R package to divide patients into high and low senescore groups in the immunotherapy and chemotherapy cohort. The independent prognostic factor was ascertained through a multivariable Cox regression model. The forest plot was employed to visualize the results of multivariate prognostic analysis using the survminer R package. All heatmaps were generated by the function of the pheatmap R package. The P values were two-side, with less than 0.05 as statistically significance. All statistical analyses were conducted in R language software (version 3.6.1).

Declarations

Acknowledgements

We sincerely thank Gene-Expression Omnibus (GEO) and the Cancer Genome Atlas (TCGA) database for data collection and supply.

Authors' contributions

Q.S., H.H. and X.W. designed the study and revised the manuscript. L.Z., Q.S. performed the data analysis and wrote the manuscript with inputs from G.M.. Z.N. and Y.W conducted the multiplex immunofluorescence immunohistochemistry and imaging. Y.Z., Y.Z., C.W., X.G. and L.G. did data collection. W.Z. and K.Z assisted the multiplex immunofluorescence immunohistochemistry imaging and analysis. All authors have read and approved the final manuscript.

Funding

This work was supported by Beijing Municipal Natural Science Foundation (KZ202110025029 to HH), Beijing Municipal Administration of Hospitals Incubating Program (PX2021033 to HH), and the National Natural Science Foundation of China (31970685 to QS, 31770975 to XW).

Data availability

The data that support findings of this study are available in the Gene Expression Omnibus (GEO, <https://www.ncbi.nlm.nih.gov/geo/>) under accession numbers GSE26253, GSE26899, GSE26901, GSE28541, GSE13861, GSE66229, GSE15459, GSE34942, GSE84437, GSE57303 and GSE29272, as well as The Cancer Genome Atlas (TCGA, <https://portal.gdc.cancer.gov>) database.

Competing interests

The authors declare that they have no competing interests.

References

1. Smyth EC, Nilsson M, Grabsch HI, van Grieken NC & Lordick F. Gastric cancer. *Lancet* **396**, 635-648 (2020).
2. Jiang Y, Li T, Liang X, Hu Y, Huang L, Liao Z *et al.* Association of Adjuvant Chemotherapy With Survival in Patients With Stage II or III Gastric Cancer. *JAMA Surgery* **152**, e171087-e171087 (2017).
3. Liu Y, Sethi NS, Hinoue T, Schneider BG, Cherniack AD, Sanchez-Vega F *et al.* Comparative Molecular Analysis of Gastrointestinal Adenocarcinomas. *Cancer Cell* **33**, 721-735 e728 (2018).
4. Cristescu R, Lee J, Nebozhyn M, Kim K-M, Ting JC, Wong SS *et al.* Molecular analysis of gastric cancer identifies subtypes associated with distinct clinical outcomes. *Nature medicine* **21**, 449-456 (2015).
5. Galluzzi L, Chan TA, Kroemer G, Wolchok JD & López-Soto A. The hallmarks of successful anticancer immunotherapy. *Science translational medicine* **10** (2018).
6. Hegde PS & Chen DS. Top 10 Challenges in Cancer Immunotherapy. *Immunity* **52**, 17-35 (2020).
7. Aaes TL & Vandenabeele P. The intrinsic immunogenic properties of cancer cell lines, immunogenic cell death, and how these influence host antitumor immune responses. *Cell Death Differ* **28**, 843-860

(2021).

8. Gorgoulis V, Adams PD, Alimonti A, Bennett DC, Bischof O, Bishop C *et al.* Cellular Senescence: Defining a Path Forward. *Cell* **179**, 813-827 (2019).
9. Shahbandi A, Rao SG, Anderson AY, Frey WD, Olayiwola JO, Ungerleider NA *et al.* BH3 mimetics selectively eliminate chemotherapy-induced senescent cells and improve response in TP53 wild-type breast cancer. *Cell Death Differ* **27**, 3097-3116 (2020).
10. Perez-Mancera PA, Young AR & Narita M. Inside and out: the activities of senescence in cancer. *Nat Rev Cancer* **14**, 547-558 (2014).
11. Xue W, Zender L, Miething C, Dickins RA, Hernando E, Krizhanovsky V *et al.* Senescence and tumour clearance is triggered by p53 restoration in murine liver carcinomas. *Nature* **445**, 656-660 (2007).
12. Arreal L, Piva M, Fernández S, Revandkar A, Schaub-Clerigué A, Villanueva J *et al.* Targeting PML in triple negative breast cancer elicits growth suppression and senescence. *Cell Death Differ* **27**, 1186-1199 (2020).
13. Coppe J-P, Patil CK, Rodier F, Sun Y, Öz DPM, Goldstein J *et al.* Senescence-Associated Secretory Phenotypes Reveal Cell-Nonautonomous Functions of Oncogenic RAS and the p53 Tumor Suppressor. *PLoS Biology* **6**, e301 (2008).
14. Jackson JG, Pant V, Li Q, Chang LL, Quintas-Cardama A, Garza D *et al.* p53-mediated senescence impairs the apoptotic response to chemotherapy and clinical outcome in breast cancer. *Cancer Cell* **21**, 793-806 (2012).
15. Xu C, Xie N, Su Y, Sun Z, Liang Y, Zhang N *et al.* HnRNP F/H associate with hTERC and telomerase holoenzyme to modulate telomerase function and promote cell proliferation. *Cell Death Differ* **27**, 1998-2013 (2020).
16. Acosta JC, Banito A, Wuestefeld T, Georgilis A, Janich P, Morton JP *et al.* A complex secretory program orchestrated by the inflammasome controls paracrine senescence. *Nat Cell Biol* **15**, 978-990 (2013).
17. Kuilman T, Michaloglou C, Vredeveld LC, Douma S, van Doorn R, Desmet CJ *et al.* Oncogene-induced senescence relayed by an interleukin-dependent inflammatory network. *Cell* **133**, 1019-1031 (2008).
18. Liu J-Y, Souroullas GP, Diekman BO, Krishnamurthy J, Hall BM, Sorrentino JA *et al.* Cells exhibiting strong promoter activation in vivo display features of senescence. *Proceedings of the National Academy of Sciences of the United States of America* **116**, 2603-2611 (2019).
19. Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP & Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst* **1**, 417-425 (2015).
20. Jung SH, Lee M, Park HA, Lee HC, Kang D, Hwang HJ *et al.* Integrin $\alpha 6 \beta 4$ -Src-AKT signaling induces cellular senescence by counteracting apoptosis in irradiated tumor cells and tissues. *Cell Death Differ* **26**, 245-259 (2019).
21. Mariathasan S, Turley SJ, Nickles D, Castiglioni A, Yuen K, Wang Y *et al.* TGF β attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. *Nature* **554** (2018).

22. Jung H, Kim HS, Kim JY, Sun J-M, Ahn JS, Ahn M-J *et al.* DNA methylation loss promotes immune evasion of tumours with high mutation and copy number load. *Nature communications* **10**, 4278 (2019).
23. Riaz N, Havel JJ, Makarov V, Desrichard A, Urba WJ, Sims JS *et al.* Tumor and Microenvironment Evolution during Immunotherapy with Nivolumab. *Cell* **171** (2017).
24. Di Micco R, Krizhanovsky V, Baker D & d'Adda di Fagagna F. Cellular senescence in ageing: from mechanisms to therapeutic opportunities. *Nat Rev Mol Cell Biol* **22**, 75-95 (2021).
25. Huang S, Li Y, Yuan X, Zhao M, Wang J, Li Y *et al.* The UbL-UBA Ubiquilin4 protein functions as a tumor suppressor in gastric cancer by p53-dependent and p53-independent regulation of p21. *Cell Death Differ* **26**, 516-530 (2019).
26. Kim ST, Cristescu R, Bass AJ, Kim K-M, Odegaard JI, Kim K *et al.* Comprehensive molecular characterization of clinical responses to PD-1 inhibition in metastatic gastric cancer. *Nature medicine* **24**, 1449-1458 (2018).
27. Panda A, Mehnert JM, Hirshfield KM, Riedlinger G, Damare S, Saunders T *et al.* Immune Activation and Benefit From Avelumab in EBV-Positive Gastric Cancer. *Journal of the National Cancer Institute* **110**, 316-320 (2018).
28. Le DT, Durham JN, Smith KN, Wang H, Bartlett BR, Aulakh LK *et al.* Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. *Science* **357**, 409-413 (2017).
29. Demaria M, O'Leary MN, Chang J, Shao L, Liu S, Alimirah F *et al.* Cellular Senescence Promotes Adverse Effects of Chemotherapy and Cancer Relapse. *Cancer Discovery* **7**, 165-176 (2016).
30. Saleh T, Tyutyunyk-Massey L & Gewirtz DA. Tumor Cell Escape from Therapy-Induced Senescence as a Model of Disease Recurrence after Dormancy. *Cancer Res* **79** (2019).
31. Mantovani F, Collavin L & Del Sal G. Mutant p53 as a guardian of the cancer cell. *Cell Death Differ* **26**, 199-212 (2019).
32. Colaprico A, Silva TC, Olsen C, Garofano L, Cava C, Garolini D *et al.* TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res* **44**, e71 (2016).
33. Davis S & Meltzer PS. GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. *Bioinformatics* **23** (2007).
34. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* **43**, e47 (2015).
35. Oh SC, Sohn BH, Cheong J-H, Kim S-B, Lee JE, Park KC *et al.* Clinical and genomic landscape of gastric cancer with a mesenchymal phenotype. *Nature communications* **9**, 1777 (2018).
36. Law CW, Chen Y, Shi W & Smyth GK. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol* **15**, R29 (2014).
37. Robinson MD, McCarthy DJ & Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140 (2010).

38. Yu G, Wang LG, Han Y & He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284-287 (2012).
39. Ellrott K, Bailey MH, Saksena G, Covington KR, Kandoth C, Stewart C *et al.* Scalable Open Science Approach for Mutation Calling of Tumor Exomes Using Multiple Genomic Pipelines. *Cell Syst* **6** (2018).
40. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO *et al.* Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Science signaling* **6**, pl1 (2013).
41. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA *et al.* The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer discovery* **2**, 401-404 (2012).
42. Mayakonda AA-OX, Lin DC, Assenov Y, Plass C & Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome research* **28**, 1747–1756 (2018).
43. Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Ou Yang TH *et al.* The Immune Landscape of Cancer. *Immunity* **48**, 812-830 e814 (2018).
44. Zhang P, Yang M, Zhang Y, Xiao S, Lai X, Tan A *et al.* Dissecting the Single-Cell Transcriptome Network Underlying Gastric Premalignant Lesions and Early Gastric Cancer. *Cell Reports* **27**, 1934-1947.e1935 (2019).
45. Zhou L, Niu Z, Jiang X, Zhang Z, Zheng Y, Wang Z *et al.* SARS-CoV-2 Targets by the pscRNA Profiling of ACE2, TMPRSS2 and Furin Proteases. *iScience* **23**, 101744 (2020).
46. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM *et al.* Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902.e1821 (2019).
47. Charrad M, Ghazzali N, Boiteau V & Niknafs A. NbClust: An R Package for Determining the Relevant Number of Clusters in a Data Set. *Journal of Statistical Software* **61** (2014).
48. Charoentong P, Francesca Finotello, Mihaela Angelova, Clemens Mayer, Mirjana Efremova, Dietmar Rieder *et al.* Pan-cancer Immunogenomic Analyses Reveal Genotype-Immunophenotype Relationships and Predictors of Response to Checkpoint Blockade. *Cell Reports* **18**, 248-262 (2017).
49. Şenbabaoğlu Y, Gejman RS, Winer AG, Liu M, Van Allen EM, de Velasco G *et al.* Tumor immune microenvironment characterization in clear cell renal cell carcinoma identifies prognostic and immunotherapeutically relevant messenger RNA signatures. *Genome biology* **17**, 231 (2016).
50. Angelova M, Charoentong P, Hackl H, Fischer ML, Snajder R, Krogsdam AM *et al.* Characterization of the immunophenotypes and antigenomes of colorectal cancers reveals distinct tumor escape mechanisms and novel targets for immunotherapy. *Genome biology* **16**, 64 (2015).
51. Hanzelmann S, Castelo R & Guinney J. GSEA: gene set variation analysis for microarray and RNA-Seq data. *BMC Bioinformatics* **14:7** (2013).
52. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545-15550 (2005).

53. Schaaf MB, Houbaert D, Meçe O & Agostinis P. Autophagy in endothelial cells and tumor angiogenesis. *Cell Death Differ* **26**, 665-679 (2019).
54. Goeman JJ. L1 Penalized Estimation in the Cox Proportional Hazards Model. *Biometrical Journal* **52**, 70-84 (2010).
55. Chen H-Y, Yu S-L, Chen C-H, Chang G-C, Chen C-Y, Yuan A *et al.* A five-gene signature and clinical outcome in non-small-cell lung cancer. *The New England journal of medicine* **356**, 11-20 (2007).
56. Sohn BH, Hwang J-E, Jang H-J, Lee H-S, Oh SC, Shim J-J *et al.* Clinical Significance of Four Molecular Subtypes of Gastric Cancer Identified by The Cancer Genome Atlas Project. *Clinical cancer research* (2017).
57. Huang H, Chen A, Wang T, Wang M, Ning X, He M *et al.* Detecting cell-in-cell structures in human tumor samples by E-cadherin/CD68/CD45 triple staining. *Oncotarget* **6**, 20278-20287 (2015).
58. Huang H, He M, Zhang Y, Zhang B, Niu Z, Zheng Y *et al.* Identification and validation of heterotypic cell-in-cell structure as an adverse prognostic predictor for young patients of resectable pancreatic ductal adenocarcinoma. *Signal Transduct Target Ther* **5**, 246-248 (2020).
59. Cai W-Y, Dong Z-N, Fu X-T, Lin L-Y, Wang L, Ye G-D *et al.* Identification of a Tumor Microenvironment-relevant Gene set-based Prognostic Signature and Related Therapy Targets in Gastric Cancer. *Theranostics* **10**, 8633-8647 (2020).

Figures

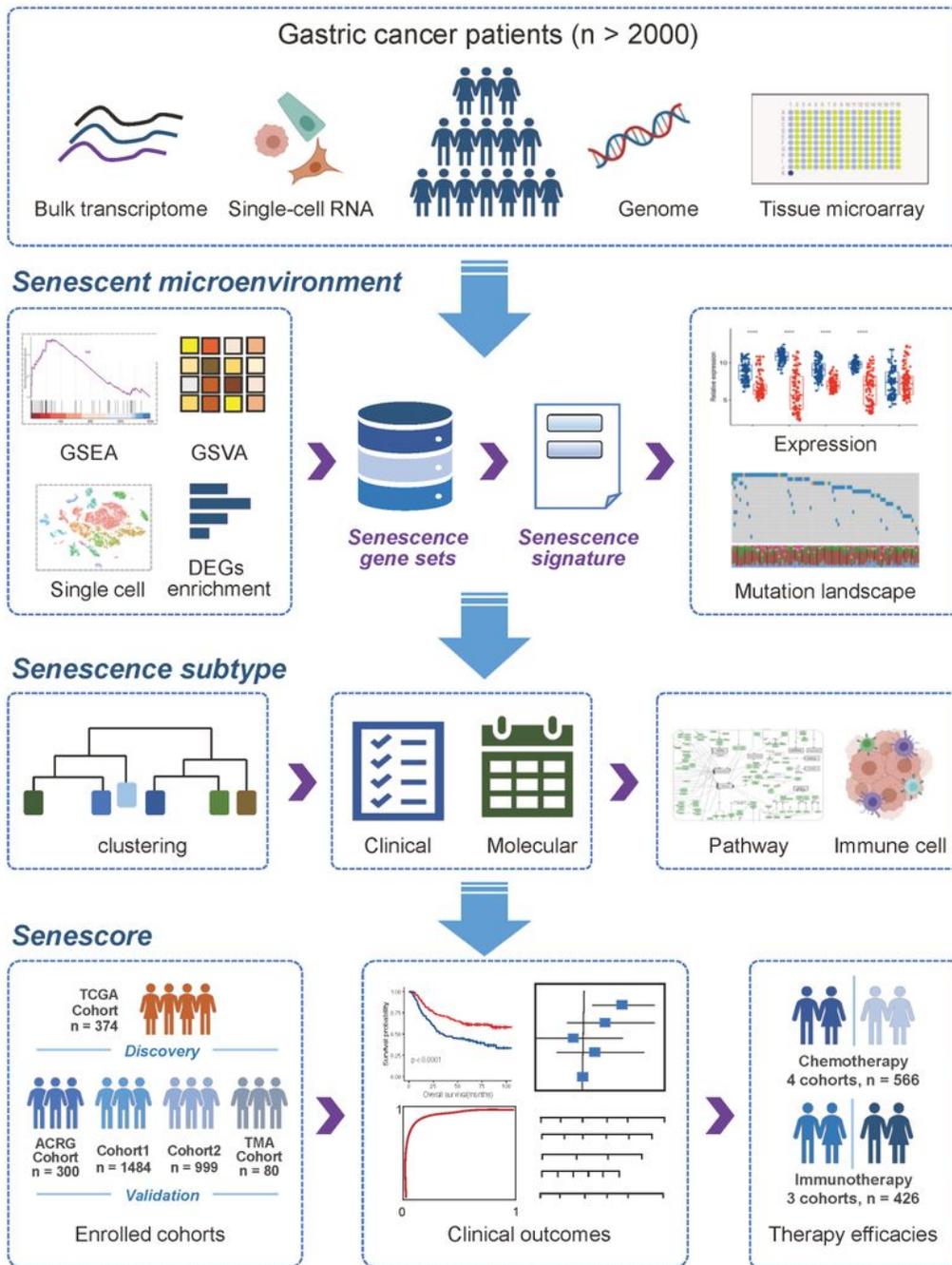


Figure 1

Figure 1

Overview of the work flow in this study. Multi-omics data ranging from bulk and single-cell transcriptome, genome to protein expression derived from more than 2000 gastric cancer patients were included in this pipeline analysis. Multiple bioinformatics algorithms including GSEA, GSVA and differential expression analysis were applied to dissect senescent microenvironment and identify the senescence signature in gastric cancer. Moreover, the senescence subtype was established through the unsupervised clustering,

and further was parsed in clinical and biological relevance. Finally, the senescence scoring system, senescore, was developed and validated for clinical prognosis and therapy efficacies in wide spectrum of patient cohorts.

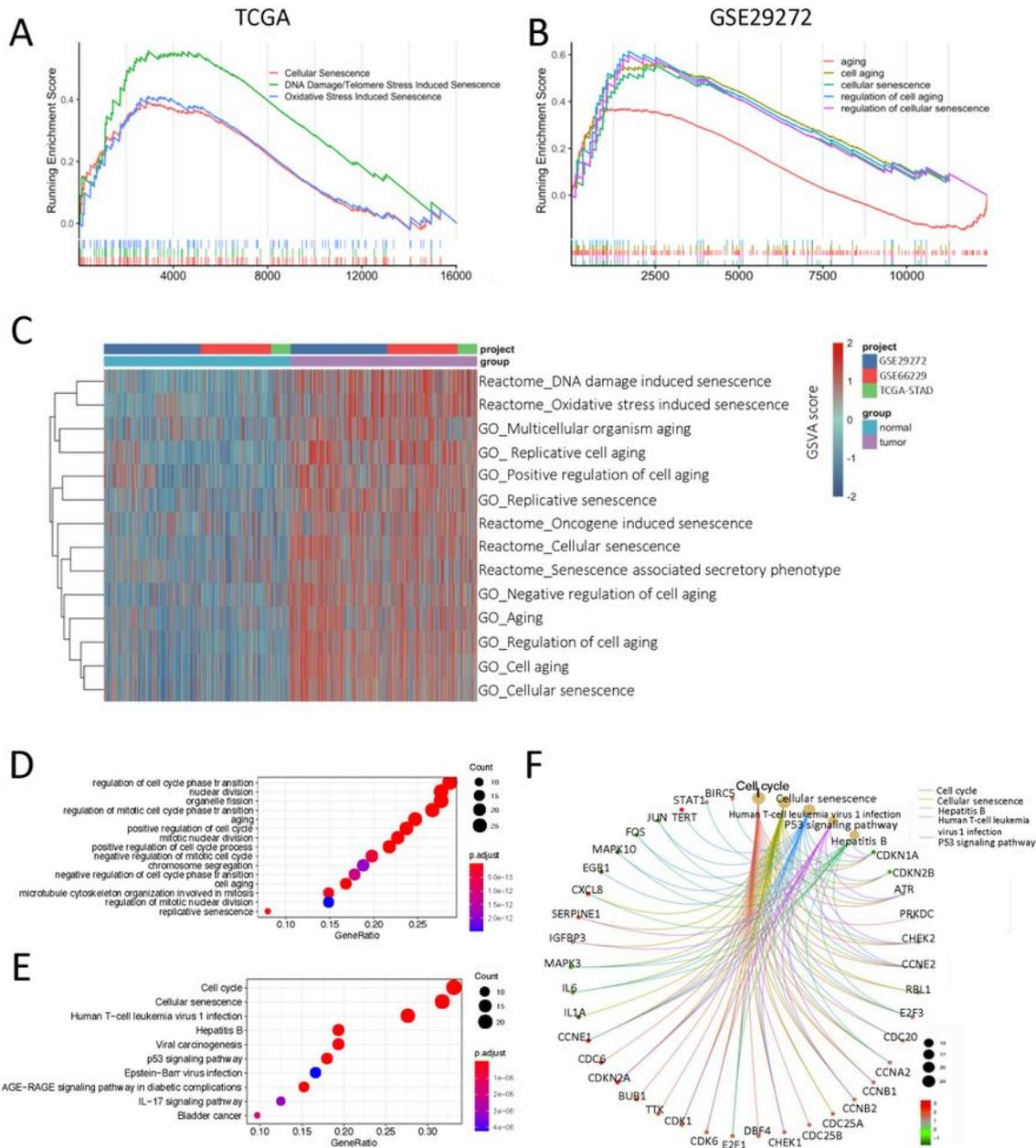


Figure 2

Figure 2

Identification of the senescence character in gastric cancer. A. GSEA plot showed that senescence-associated pathways were significantly enriched in GC relative to normal tissues in TCGA-STAD cohort.

Multiple pathway enrichment results were indicated by corresponding colors. B. GSEA plot exhibited the enriched GO terms associated with senescence based on GC versus normal gastric tissues in GSE29272 cohort. The different GO terms were displayed by corresponding colors. C. GSVA enrichment analysis showing the quite different senescence features between GC and their paired adjacent normal tissues in multiple datasets. Heatmap displaying the GSVA score of senescence pathways and GO terms. D-E. Dot plots showing the significant enrichment results for senescence-associated genes in GO term (D, top 15) and pathway (E, top 10). The dot size and color represented the gene count and the weighted Fisher's p-value, respectively. F. The interaction network exhibiting the top five enriched pathways and their contained genes. The dot size of the pathways represented the gene count. The color of genes indicated the fold change of gene expression between GC and normal tissues in TCGA-STAD cohort.

showed that the senescence core genes were established by overlapping among the DEGs, senescence genes and survival genes. F. The interaction among CSGs in GC. The risk genes and protective genes were marked with red and green, respectively. The circle size represented p value of univariate cox regression analysis. The lines connecting genes showed their interactions based on correlation. Positive correlation was marked with red and negative correlation with gray. G. Dot plot showed the expression of CSGs of each cell type in GC H. The mutation landscape of CSGs in patients from TCGA-STAD cohort. Each column represented individual patients. The upper barplot showed TMB, and the number on the right presented the mutation frequency in each gene. The barplot in right showed the proportion of each variant type. The stacked barplot below showed fraction of conversions in each sample. PMC: pit mucous cells; GMC: gland mucous cells; SMC: smooth muscle cells; DEGs: differently expressed genes.

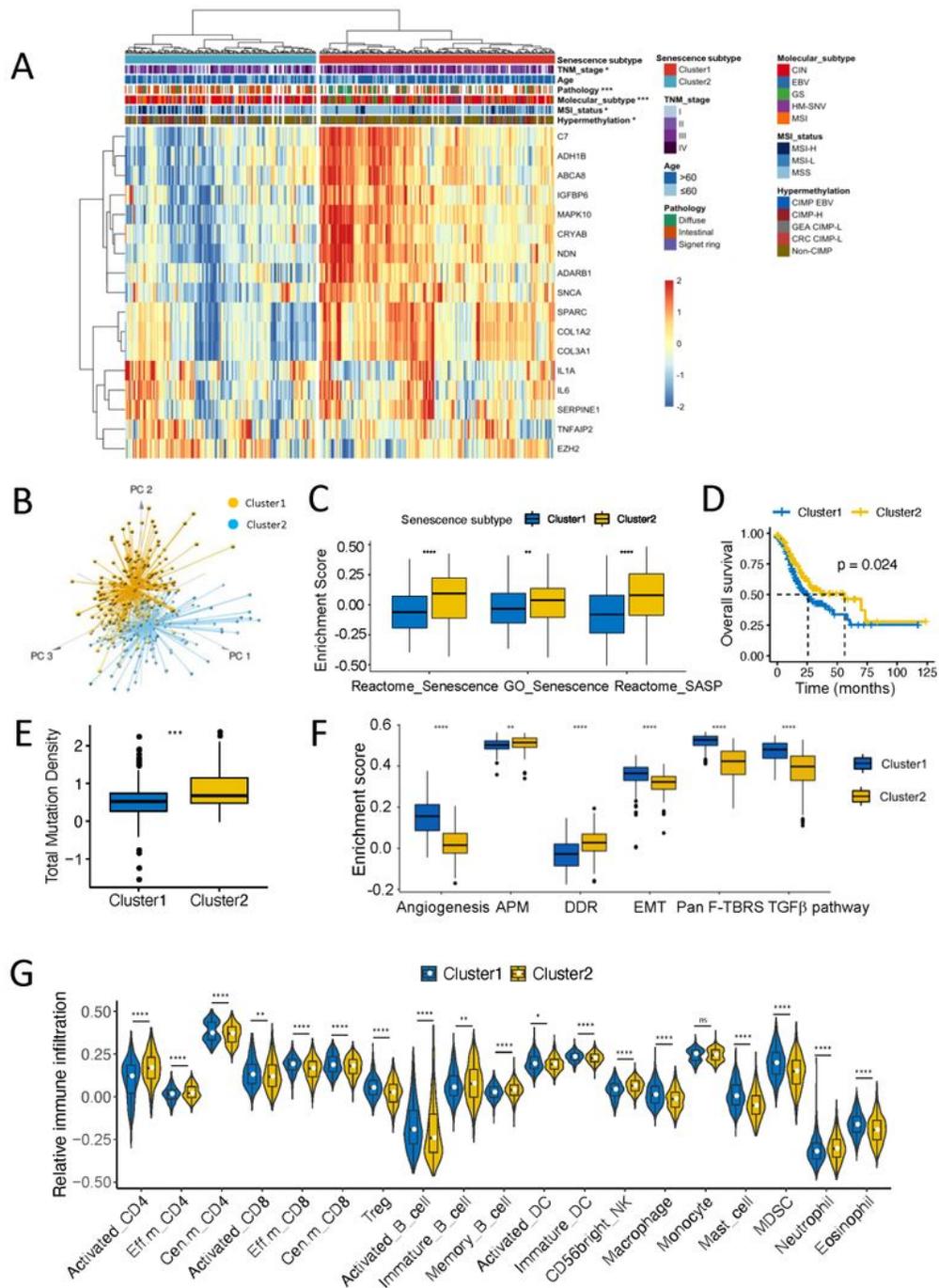


Figure 4

Figure 4

Development of senescence subtypes based on senescence core genes. A. Unsupervised clustering of patients derived from the TCGA-STAD cohort based on CSGs to classify patients into two senescence subtypes. TNM stage, age pathological classification, molecular subtype, MSI status and hypermethylation category were shown above the heatmap. B. Principal component analysis for the expression profiles to distinguish two clusters in TCGA-STAD cohort. Cluster 1 was marked with yellow

and cluster 2 was marked with blue. C. GSVA analysis for senescence gene sets indicating the significant distinction of senescence features in two subtypes. D. Survival analyses for the two subtypes in TCGA-STAD cohort. Kaplan-Meier curves showed a significant survival difference between two clusters. P value was calculated by Log-rank test. E. The significant distinction in total mutation density between senescence subtypes. F. Differences in pathway activation including angiogenesis, APM, DDR, EMT, Pan-F-TBRS and TGF β pathway between two subtypes. In C, E and F, the box represents the interquartile range, the horizontal line in the box is the median, and the whiskers represent 1.5 times the interquartile range. G. Violin plot depicting the abundance of each cell type in different clusters in the validation cohort1. In C, E, F and G, the asterisks represented the statistical p value (Wilcoxon's rank sum test, *P < 0.05; **P < 0.01; ***P < 0.001). MSI: microsatellite instability; HM-SNV: hypermutated-single-nucleotide variants; CIMP: CpG island methylator phenotype; MSS: microsatellite stable; SASP: senescence-associated secretory phenotype; APM: antigen processing machinery; DDR: DNA damage repair response; EMT: epithelial mesenchymal transition; Pan F-TBRS: pan-fibroblast TGF β response signature; Eff.m: effector memory; Cen.m: Central memory, Treg: regulatory T cells; DC: dendritic cells; NK: natural killer; MDSC: myeloid-derived suppressor cells.

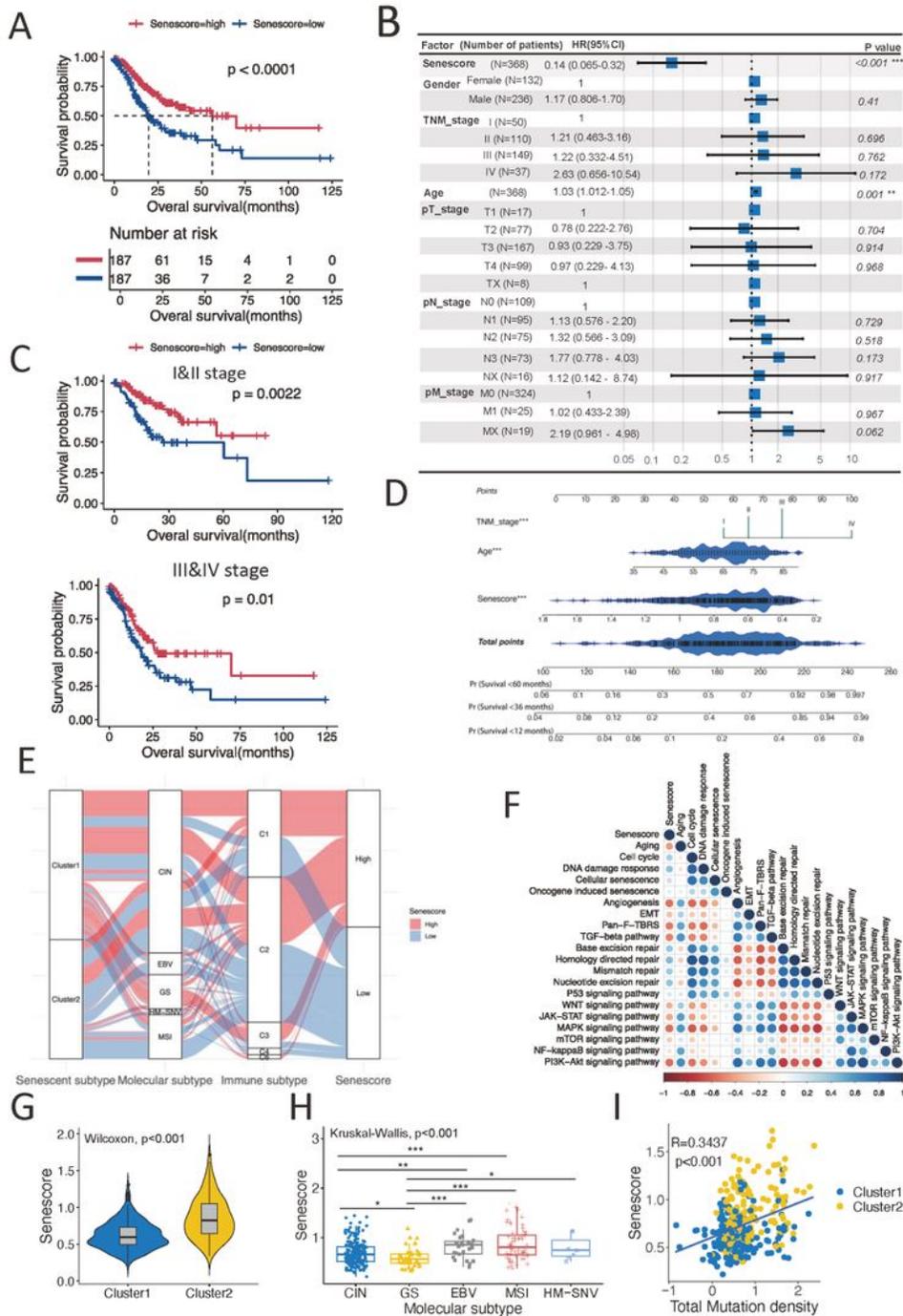


Figure 5

Figure 5

Construction of the senescence score with clinical utility. A. Kaplan-Meier curves indicated the significant survival difference between high and low senescore groups in TCGA-STAD cohort. Log-rank test, $P < 0.0001$. B. Multivariate Cox regression analysis for senescore and other clinical characteristics with OS in TCGA-STAD cohort shown by the forest plot. C. Kaplan-Meier curves showing patients with stage I/II (upper panel) or stage III/IV (lower panel) was further grouped into different survival by senescore. Log-

rank test, p value was indicated. D. Nomogram was developed in the TCGA-STAD cohort based on the senescore, age and TNM stage to predict 1- year, 3- year, and 5- year outcomes. E. Alluvial diagram showing the mutuality of the senescence subtypes, molecular subtypes, immune subtypes and senescore in TCGA-STAD cohort. F. Correlations between senescore and myriad biological processes in TCGA-STAD cohort using Spearman analysis. Negative correlation was marked with red and positive correlation with blue. The size of dots represents the value of correlation coefficient. G. Violin plot displaying the differences of senescore between senescence subtypes in TCGA-STAD cohort. The Wilcoxon rank-sum test was used to compare the statistical significance between two groups. H. The box plot showing the distinctions of the senescore in different molecular subtypes. The asterisks represented the statistical p value (the Wilcoxon rank-sum test, *P < 0.05; **P < 0.01; ***P < 0.001). The Kruskal-Wallis test was applied to evaluate significant differences among all five variables. I. Scatterplot showing the correlations of the senescore with mutation density. R coefficient represents of Pearson's correlation. P value was from the Pearson correlation test. The shaded area represents 95% confident interval. The blue and yellow colors of the dots respectively represent the cluster 1 and cluster 2.

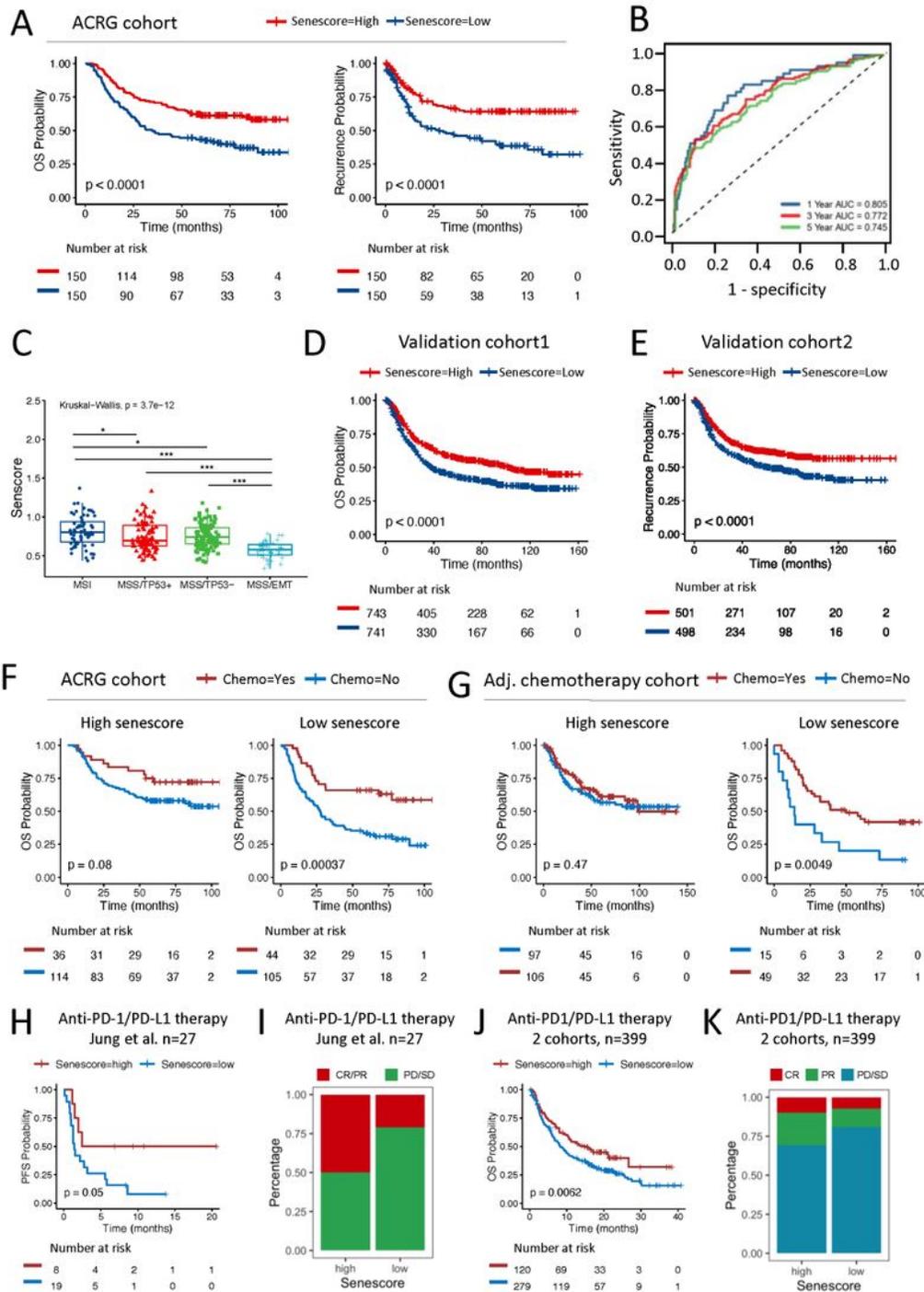


Figure 6

Figure 6

Validation of the senescore for clinical prognosis and therapeutic efficacy. A. Kaplan–Meier plots showing the differences in OS (left) and RFS (right) between the high and low senescore groups in the ACRG cohort. P value was calculated by Log-rank test. B. Time-dependent ROC curve for the validation of nomogram in the ACRG cohort. C. Differences of the senescore among molecular subtypes in the ACRG cohort (the Wilcoxon rank-sum test, *P < 0.05; **P < 0.01; ***P < 0.001). The Kruskal-Wallis test was

applied to evaluate significant differences among four variables. D. Kaplan–Meier plot indicating the distinctions in patients' OS between high and low senescore groups in the validation cohort1. E. Kaplan–Meier plot displaying the predictive value of the senescore for patients' RFS in the validation cohort2. F. Kaplan–Meier plots of GC patients of the high and low senescore group in ACRG cohort. The data in each group were analyzed according to whether patients received adjuvant chemotherapy or not. G. Survival analyses for patients in adjuvant chemotherapy cohort were stratified by both the senescore and treatment with adjuvant chemotherapy using Kaplan-Meier curves. H. Survival analyses for low and high senescore patients in an independent anti-PD-L1 therapy cohort using Kaplan-Meier curves. I. The proportion of patients with response to PD1/PD-L1 blockade immunotherapy in the patients with high or low senescore. J. Kaplan-Meier curves showed survival differences of OS between low and high senescore patients in the integrated cohort of anti-PD1/PD-L1 therapy. K. The proportion of patients with the therapy response to PD1/PD-L1 blockade immunotherapy in the high or low senescore groups. In D, E, F, G, H, and J, p value was calculated by Log-rank test. OS: overall survival; RFS: recurrence-free survival; Chemo: chemotherapy; PFS: progress free survival; CR: complete response; PR: partial response; PD: progressive disease; SD: stable disease.

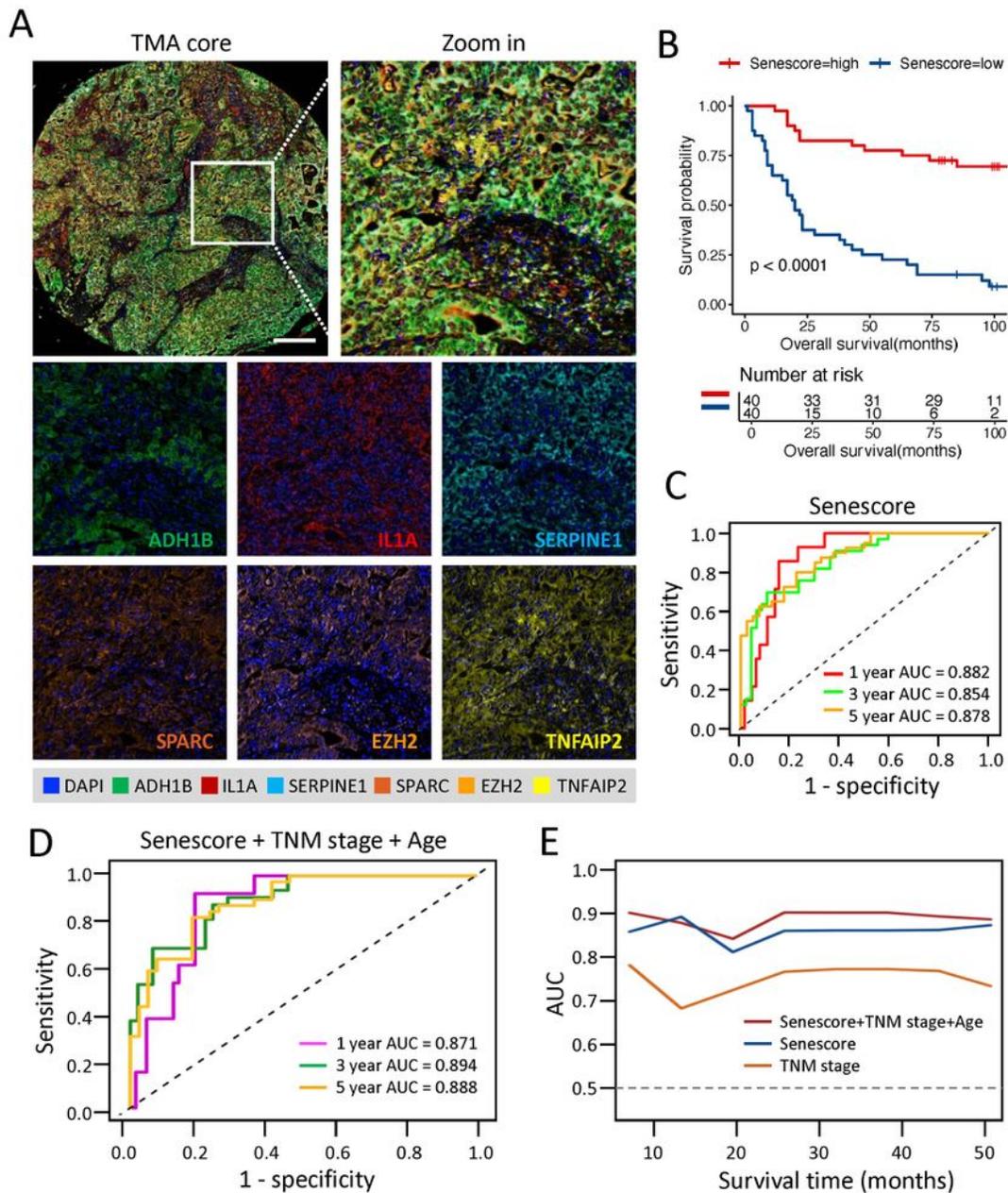


Figure 7

Figure 7

Multiplex immunofluorescence immunohistochemistry profiling of the senescore in tissue microarray of GC. A. Representative multiplex immunofluorescence immunohistochemistry images showed the protein expression patterns of ADH1B (green), IL1A (red), SERPINE1(cyan), SPARC (orange), EZH2 (pink) and TNFAIP2 (yellow) in GC tissues. DAPI: blue; scale bar: 200 μ m. B. Kaplan-Meier plot showed the significant difference of OS between high and low senescore groups in the tissue microarrays of GC. Log-rank test, p

value was indicated. C. ROC curve showed the predictive capability of senescore for overall survival of GC patients. D. ROC curves of the senescore combined with clinical stage and age for predicting 1-year, 3-year, and 5-year survival. E. Comparison of AUC of the time-dependent ROC curve among nomogram containing senescore, clinical stage and age, and only senescore or clinical stages.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Figuresupplement.pdf](#)
- [SupplementaryFigureLegends.docx](#)
- [TableS1.xlsx](#)
- [TableS2.xlsx](#)
- [TableS3DetailedinformationofTCGASTADcohortinthisstudy.xlsx](#)
- [TableS4Detailsaboutthedatasetsusedinthisstudy.xlsx](#)
- [TableS5Detailedinformationofchemotherapycohortsinthisstudy.xlsx](#)
- [TableS6Detailedinformationofimmunotherapycohortsinthisstudy.xlsx](#)
- [TableS7Detailedinformationoftissuemicroarraycohortinthisstudy.xlsx](#)