Identification of potential shared gene signatures between gastric cancer and type 2 diabetes: a data-driven analysis

Bingqing Xia[1], Ping Zeng[1], Yuling Xue[1], Qian Li[2,3,4], Jianhui Xie[2,3,4], Jiamin Xu[2,3,4], Wenzhen Wu[2,3,4] and Xiaobo Yang[2,3,4]*

[1]The Second Clinical College of Guangzhou University of Chinese Medicine, Guangzhou, China, [2]The Second Affiliated Hospital of Guangzhou University of Chinese Medicine, Guangzhou, China, [3]State Key Laboratory of Dampness Syndrome of Chinese Medicine, The Second Affiliated Hospital of Guangzhou University of Chinese Medicine, Guangzhou, China, [4]Guangdong Provincial Key Laboratory of Clinical Research on Traditional Chinese Medicine Syndrome, Guangzhou, China

Background: Gastric cancer (GC) and type 2 diabetes (T2D) contribute to each other, but the interaction mechanisms remain undiscovered. The goal of this research was to explore shared genes as well as crosstalk mechanisms between GC and T2D.

Methods: The Gene Expression Omnibus (GEO) database served as the source of the GC and T2D datasets. The differentially expressed genes (DEGs) and weighted gene co-expression network analysis (WGCNA) were utilized to identify representative genes. In addition, overlapping genes between the representative genes of the two diseases were used for functional enrichment analysis and protein−protein interaction (PPI) network. Next, hub genes were filtered through two machine learning algorithms. Finally, external validation was undertaken with data from the Cancer Genome Atlas (TCGA) database.

Results: A total of 292 and 541 DEGs were obtained from the GC (GSE29272) and T2D (GSE164416) datasets, respectively. In addition, 2,704 and 336 module genes were identified in GC and T2D. Following their intersection, 104 crosstalk genes were identified. Enrichment analysis indicated that "ECM-receptor interaction," "AGE-RAGE signaling pathway in diabetic complications," "aging," and "cellular response to copper ion" were mutual pathways. Through the PPI network, 10 genes were identified as candidate hub genes. Machine learning further selected BGN, VCAN, FN1, FBLN1, COL4A5, COL1A1, and COL6A3 as hub genes.

Conclusion: "ECM-receptor interaction," "AGE-RAGE signaling pathway in diabetic complications," "aging," and "cellular response to copper ion" were revealed as possible crosstalk mechanisms. BGN, VCAN, FN1, FBLN1, COL4A5, COL1A1, and COL6A3 were identified as shared genes and potential therapeutic targets for people suffering from GC and T2D.

KEYWORDS

bioinformatics, gastric cancer, type 2 diabetes, crosstalk genes, pathways

1 Introduction

As a prevalent tumor worldwide, gastric cancer (GC) has a relatively terrible prognosis. In 2020, estimates from GLOBOCAN revealed that GC ranked fourth for mortality and fifth for morbidity globally (1), which dramatically increased the burden of finance and medical care (2, 3). Type 2 diabetes (T2D) is among the most prevalent endocrine diseases (4). A study representing 215 countries and regions has provided evidence that the global incidence of diabetes will increase, ranging from approximately 10.5% in 2021 to 12.2% in 2045 (5, 6).

T2D is a risk factor for the emergence of certain cancers (7) and is connected to a higher cancer death rate (8, 9). According to cohort research involving over 46,000 patients, diabetes is linked to a 67% increase in GC risk (10). It is also one of the main non-cancer contributors to death in GC (11). Independent of conventional diabetes risk variables, a large-scale cohort investigation has verified that GC is linked to an elevated risk of diabetes (7). As the risk of developing T2D increases by 35% (7) in GC, the mortality of GC survivors also significantly increases (12).

All those empirical observations suggest a strong bidirectional association underlying these two complex diseases. Though the associations between GC and T2D have been widely reported, the underlying genetic processes linking GC and T2D are still unknown. Fortunately, genetic exploration of disease–disease interaction has been possible through advances in sequencing and bioinformatics (4, 13–16). Meanwhile, with the continuous development of machine learning algorithms in medicine, numerous studies have applied them to the screening of feature biomarkers (17–19).

As is well known, exploring shared gene signatures between GC and T2D is of great significance for developing novel therapeutic strategies for joint prediction, prevention, and intervention. Therefore, the goal of our study was to uncover pivotal shared genes and related mechanisms between GC and T2D via bioinformatics and machine learning algorithms.

2 Materials and methods

2.1 Source of the dataset

"Stomach cancer" and "Type 2 diabetes" were used for digging the Gene Expression Omnibus (GEO) database[1] for datasets related to either condition. The requirements comprised: (1) *homo sapiens*; (2) each dataset had a sample size of at least 15; (3) the tested tissues were gastric tissue or pancreatic islets; and (4) one disease group and one control group should be included in the dataset.

The annotation soft tables were downloaded from the relevant GPL platform. Probe ID and ensemble ID were converted to official gene symbols through R (v4.2.0) and Perl (v5.30.3), respectively.

2.2 Differentially expressed gene screen

GC-related datasets further identified the differentially expressed genes (DEGs) using limma (20) R (version 3.5.1), while T2D-related datasets were analyzed using edge (21) R packages version 3.19. Raw datasets underwent normalization to identify potential mechanisms and relevant biological characteristics associated with pathways of DEGs in GC and T2D. Subsequently, the normalized datasets were transformed using the log2 function. When a replicated gene symbol is mapped to the identical gene, the average expression intensity is taken. DEGs need to meet $|logFC| \geq 1$ as well as an adjusted *p*-value of < 0.05. DEGs were shown as heatmaps and volcano plots through ggplot2 (22) and pheatmap (23) packages.

2.3 Weighted gene co-expression network analysis

The GC and T2D datasets were subjected to WGCNA through the WGCNA (24) package. First, the goodSamplesGenes function was used to filter qualified samples and genes from the expression matrix to produce a scale-free co-expression. Outliers in the remaining samples were detected using cluster analysis. Next, "pickSoftThreshold" function was utilized to determine the network topology information, and the appropriate β value was selected as the soft threshold to construct the network to make the network meet the scale-free topology characteristics (25). Pearson analysis was used to compute the gene correlations matrix files. Then, the adjacency matrix was constructed by combining the gene correlations matrix and β. In the next step, we first constructed a topological overlap matrix (TOM) by transforming the adjacency matrix, allowing us to assess the gene relationships and dissimilarities within the network. Subsequently, hierarchical clustering was used to group genes based on their similarity in expression profiles, and the dynamic tree-cut function was applied to identify distinct modules within the network. Additionally, we computed module eigengenes (MEs) by evaluating the correlation coefficients between each module and the gene expression patterns. This facilitated the identification of modules strongly associated with specific biological processes (BPs) or conditions within the dataset. Genes exhibiting strong connections within these modules were then selected for further investigation.

2.4 Identification of crosstalk genes

DEGs, as well as genes in related ME, were considered representative genes for the disease. Thus, crosstalk genes were defined as genes that overlapped between the GC-related and T2D-related representative genes.

2.5 Functional enrichment analysis

The top 10 pathways or BPs were selected based on their significance scores derived from the DAVID website's[2] analysis. Specifically, the Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases were utilized (26, 27). The enrichment results were then processed using the ggplot2 (22) and stringr (28) packages in the R programming language to generate visualizations. ggplot2 was used for creating graphical representations

---

1 http://www.ncbi.nlm.nih.gov/geo/

2 https://david.ncifcrf.gov/summary.jsp

of the enrichment results, while stringr facilitated data manipulation tasks, such as parsing and formatting pathway names for improved visualization clarity. This approach allowed for the clear visualization of the most relevant pathways and BPs implicated in the analyzed gene expression data.

2.6 Protein−protein interaction network

The STRING database[3] (29) yielded PPI networks for crosstalk genes. The outcomes were then fully visualized using Cytoscape software Version 3.10.2 (30). MCODE with filter criteria including a degree cutoff of 2, a node score cutoff of 0.2, a k-core of 2, and a maximum depth of 100 and cytoHubba (using MCC) plugins were used to screen core gene clusters/hub genes (31–34). Candidate hub genes were defined as the top 10 genes ranked by the cytoHubba plugin and genes in the first cluster chosen by the MCODE plugin.

2.7 Machine learning

Support vector machine-recursive feature elimination (SVM-RFE) (35) and least absolute shrinkage and selection operator (LASSO) algorithms were used to select hub genes for GC diagnosis. SVM-RFE minimizes the feature set and identifies predictive features by training a subset of features from various categories using 10-fold cross-validation (36). The SVM-RFE algorithm was implemented for hub gene selection utilizing the kernlab (37, 38), caret (39), and e1071 (40) packages. LASSO regression could increase the predictability and understandability of a statistical model through variable selection (41). LASSO regression was carried out via glmnet (42) package, with a 10-fold cross-verification as the turning/penalty parameter. Hub genes of GC diagnosis were the intersection genes of the two machine learning algorithms.

2.8 External validation

To validate hub genes, we obtained clinical and gene data of GC cases from the Cancer Genome Atlas (TCGA) database.[4] A volcano plot demonstrated differential expression of hub genes, as well as gene expression comparison between cancer groups and control groups, which was further illustrated in a boxplot using reshape2 (43) and ggpubr (44) packages. ROC curves for hub genes were constructed using pROC (45) and ggplot2 (22) packages. Furthermore, clinical data related to GC obtained through the TCGA database were used for survival analysis (46, 47) within the R v4.2.0 environment for survival analysis. Additionally, the survminer package (48) was utilized for visualization purposes. This enabled us to assess the prognostic significance of identified hub genes, in relation to overall survival rates among GC patients.

---

3 Results

3.1 Screening of differentially expressed genes

GSE29272 (GPL96 platform, expression profiling by array) and GSE164416 (GPL16791 platform, expression profiling by high-throughput sequencing) met our inclusion criteria. GSE29272 contained 134 GC samples and 134 control samples, while GSE164416 comprised 39 T2D samples and 18 control samples.

Following the investigation of differential gene expression, 292 DEGs (127 downregulated genes and 165 upregulated genes) were found in the GC dataset, and 541 DEGs (57 downregulated genes and 484 upregulated genes) were discovered in the T2D dataset. Figure 1 displays the volcano plot and heatmaps for DEGs.

3.2 Critical module recognition

Based on WGCNA, 21 modules were discovered in the GSE29272 ($\beta = 8$), which were represented in different colors. Through the Spearman correlation coefficient, the "blue" ($r = -0.73$, $p = 3e{-}45$, 1,863 genes), the "purple" ($r = -0.68$, $p = 3e{-}37$, 172 genes), and the "green" ($r = 0.67$, $p = 9e{-}36$, 669 genes) had a high association with GC. Additionally, 31 modules were identified in the GSE164416 ($\beta = 6$) with the "grey60" ($r = 0.72$, $p = 1e{-}09$, 138 genes) and the "salmon" ($r = -0.56$, $p = 1e{-}05$, 228 genes) having a high association with T2D. The WGCNA results are shown in Figure 2.
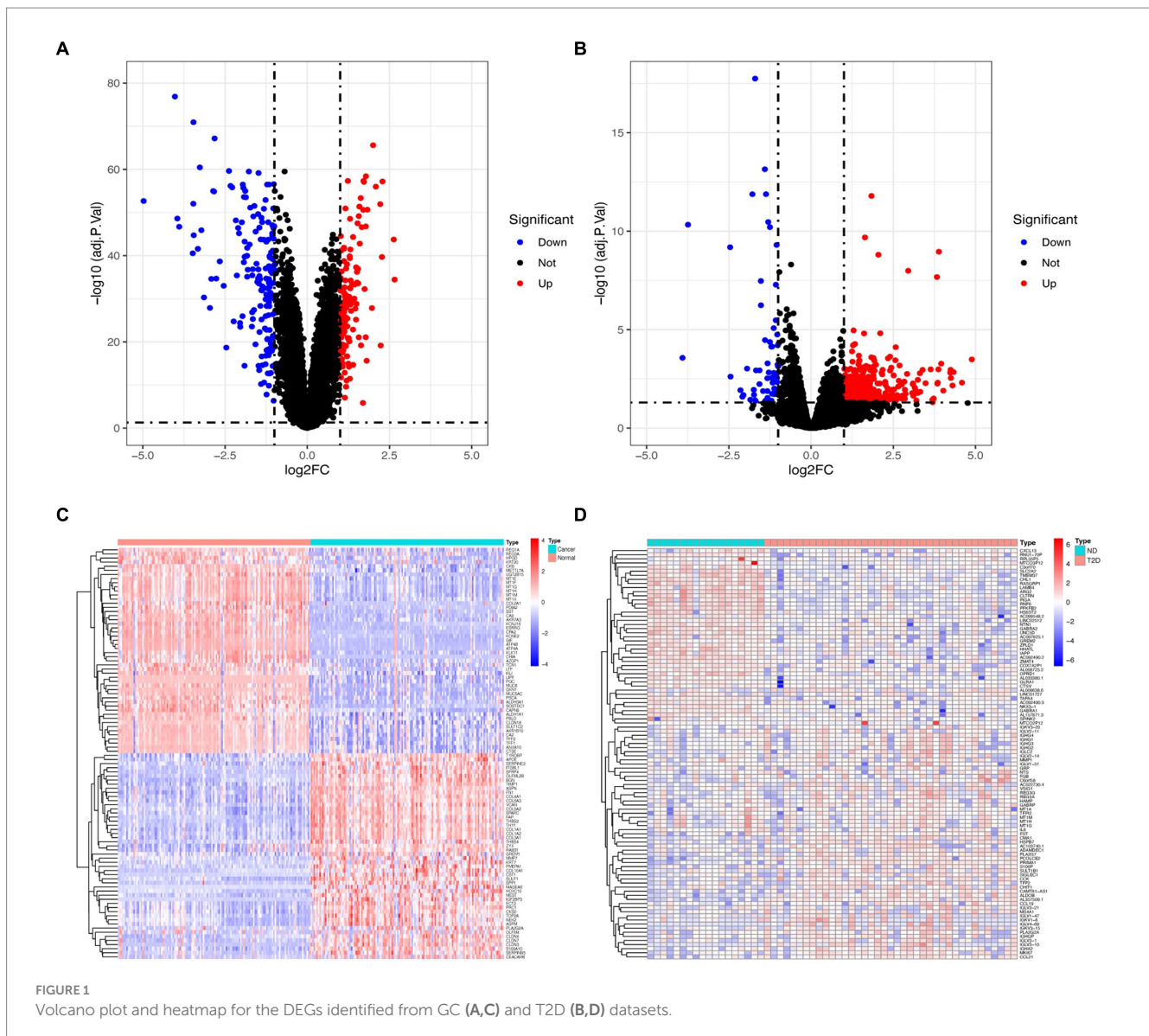
3.3 Crosstalk genes between GC and T2D

After the intersection of the representative genes between GC and T2D, 104 genes were identified as the crosstalk genes bridging GC and T2D (Figure 3A; Supplementary Table S1).

3.4 Functional enrichment analysis of crosstalk genes

The bulk of crosstalk genes was concentrated in "Protein digestion and absorption," "ECM-receptor interaction," and "AGE-RAGE signaling pathway in diabetic complications" through the KEGG analysis (Figure 3B).

In accordance with the GO analysis, the crosstalk genes were primarily enriched in BP categories, including "aging," "cellular response to copper ion," and "cerebral cortex development" (Figure 3C). As for cellular component (CC) categories, the crosstalk genes were primarily found in "extracellular region," "extracellular space," and "extracellular matrix" (Figure 3D). The three most important molecular functions (MFs) of the crosstalk genes were "extracellular matrix structural constituent conferring tensile strength," "extracellular matrix structural constituent," and "calcium ion binding" (Figure 3E). More detailed information is listed in Supplementary Table S2.

FIGURE 1
Volcano plot and heatmap for the DEGs identified from GC **(A,C)** and T2D **(B,D)** datasets.

3.5 Identification of candidate hub genes

The PPI network identified 66 interacting node genes and 120 edges in the crosstalk genes for subsequent machine learning filtration (Figure 4F). MCODE and cytoHubba plugins selected 10 genes as closely related genes, respectively. Surprisingly, the 10 genes (FN1, COL1A1, COL3A1, COL6A3, COL11A1, BGN, FBLN1, COL6A2, VCAN, and COL4A5) screened by these two plugins were the same.

3.6 Determining hub genes with machine learning

Seven candidate hub genes were selected through the LASSO regression algorithm (lambda was set to a minimum in order to minimize the regularization effect of the LASSO algorithm) (Figure 5A). In addition, nine candidate hub genes were screened through the SVM-RFE algorithm (Figure 5B). The Venn diagram

revealed that the LASSO regression and SVM-RFE algorithms identified seven cross genes (BGN, COL1A1, VCAN, FN1, COL6A3, COL4A5, and FBLN1), which were determined to be hub genes for GC diagnosis in the final validation (Figure 5C).

3.7 External validation and survival analysis

The TCGA database consisted of 412 GC samples and 36 control samples. The volcano plot (Figure 4A) and boxplot (Figure 4B) reveal that all seven genes were differently expressed between the two groups. Specifically, COL4A5 and FBLN1 were downregulated genes, while BGN, COL1A1, VCAN, FN1, and COL6A3 were upregulated. Figures 4C,D display the ROC curves corresponding to upregulated and downregulated genes separately. The AUC percentage for BGN, COL1A1, VCAN, FN1, COL6A3, COL4A5, and FBLN1 were 93, 95, 82, 70, 74, 82, and 76%, respectively. The survival analysis of the hub genes detected (COL4A5, FN1, FBLN1, and VCAN) could predict the prognosis of GC (Figures 4E–K).
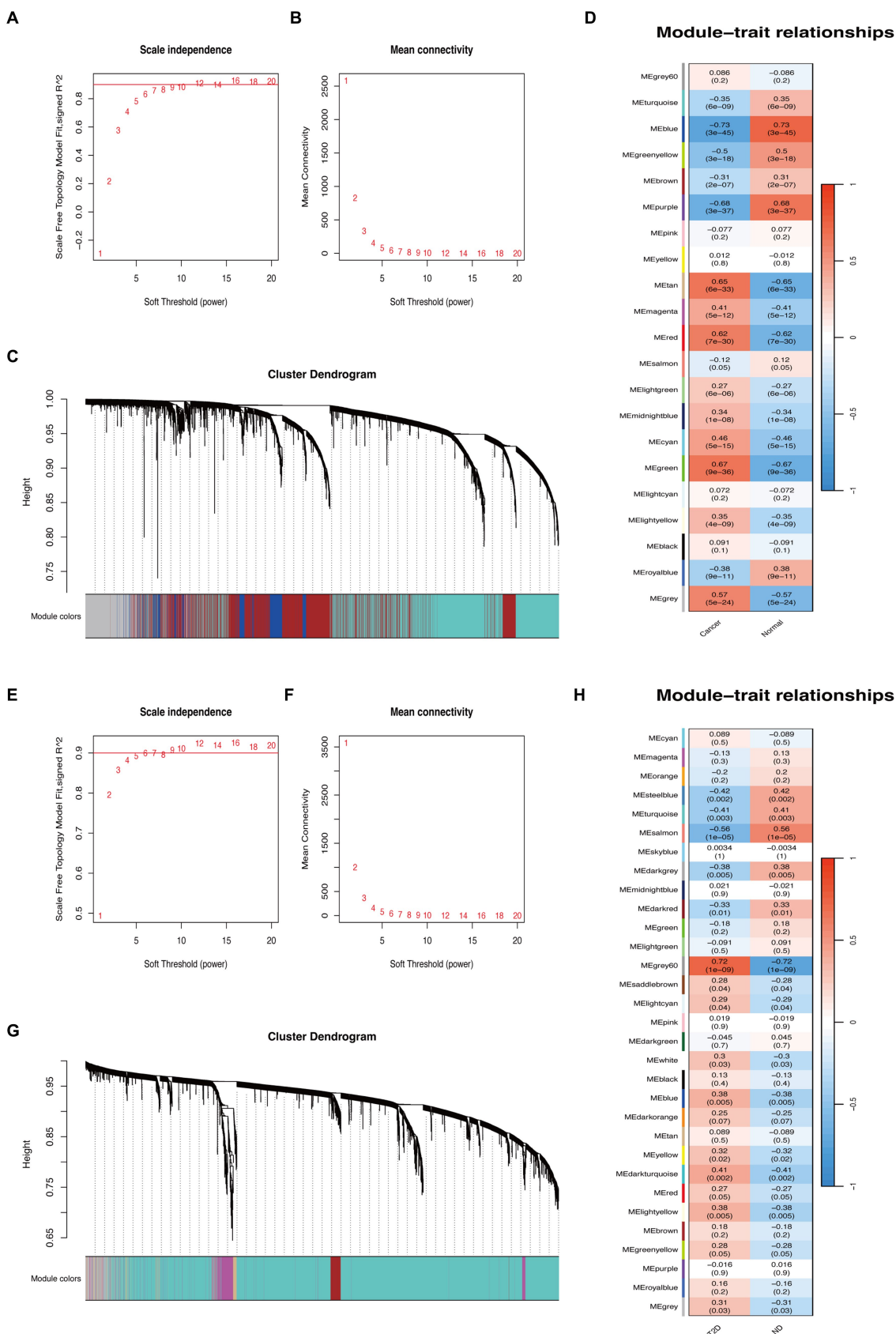
FIGURE 2
WGCNA of GC and T2D datasets. **(A,B)** β = 8 was chosen as the soft threshold for the GC dataset. **(C)** Cluster dendrogram of co-expressed genes in GC. Under the gene tree, each color represented a module. **(D)** Module−trait relationships in the GC dataset. **(E,F)** β = 6 was chosen as the soft threshold for the T2D dataset. **(G)** Cluster dendrogram of co-expressed genes in T2D. Under the gene tree, each color represented a module. **(H)** Module−trait relationships in the T2D dataset.
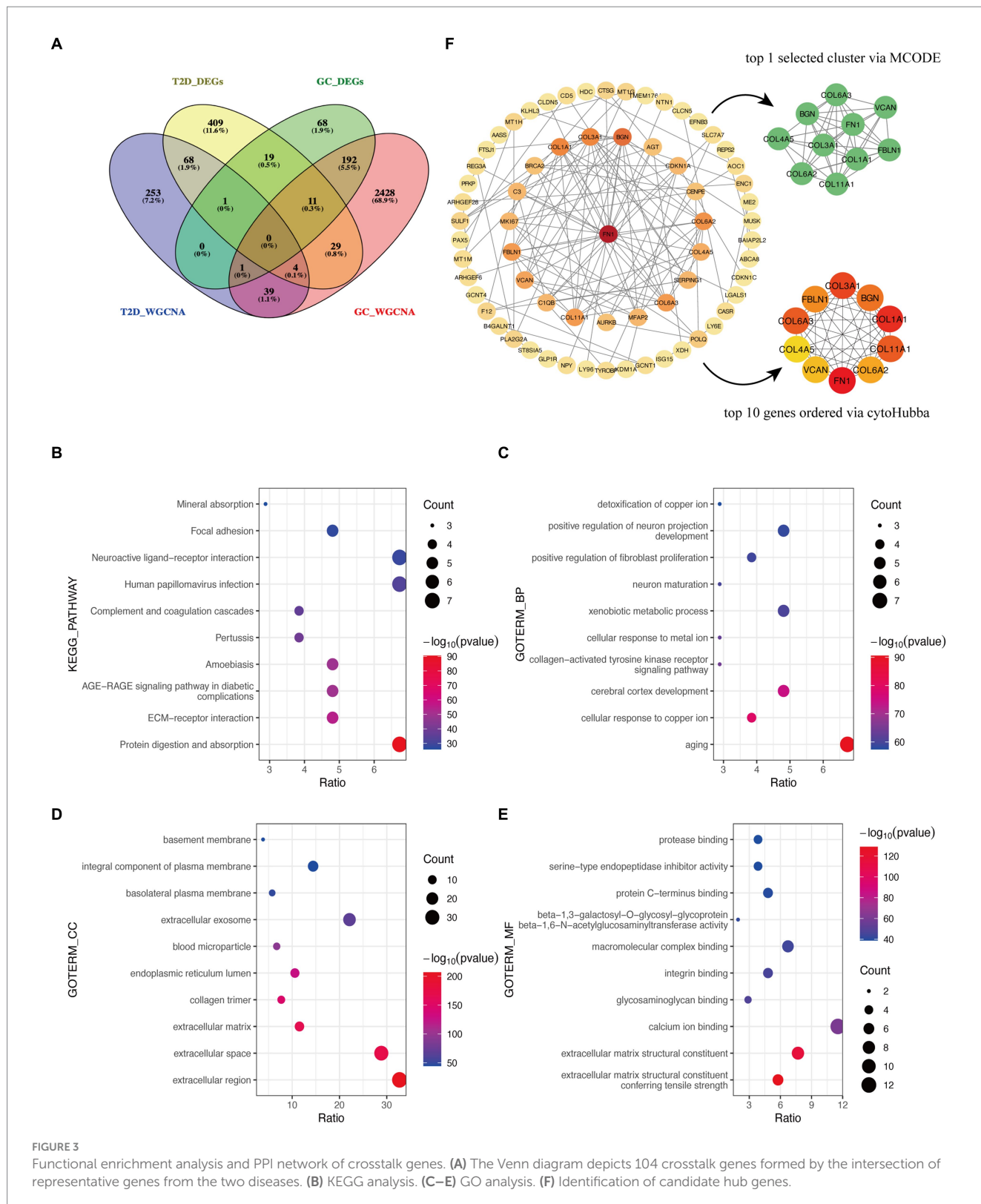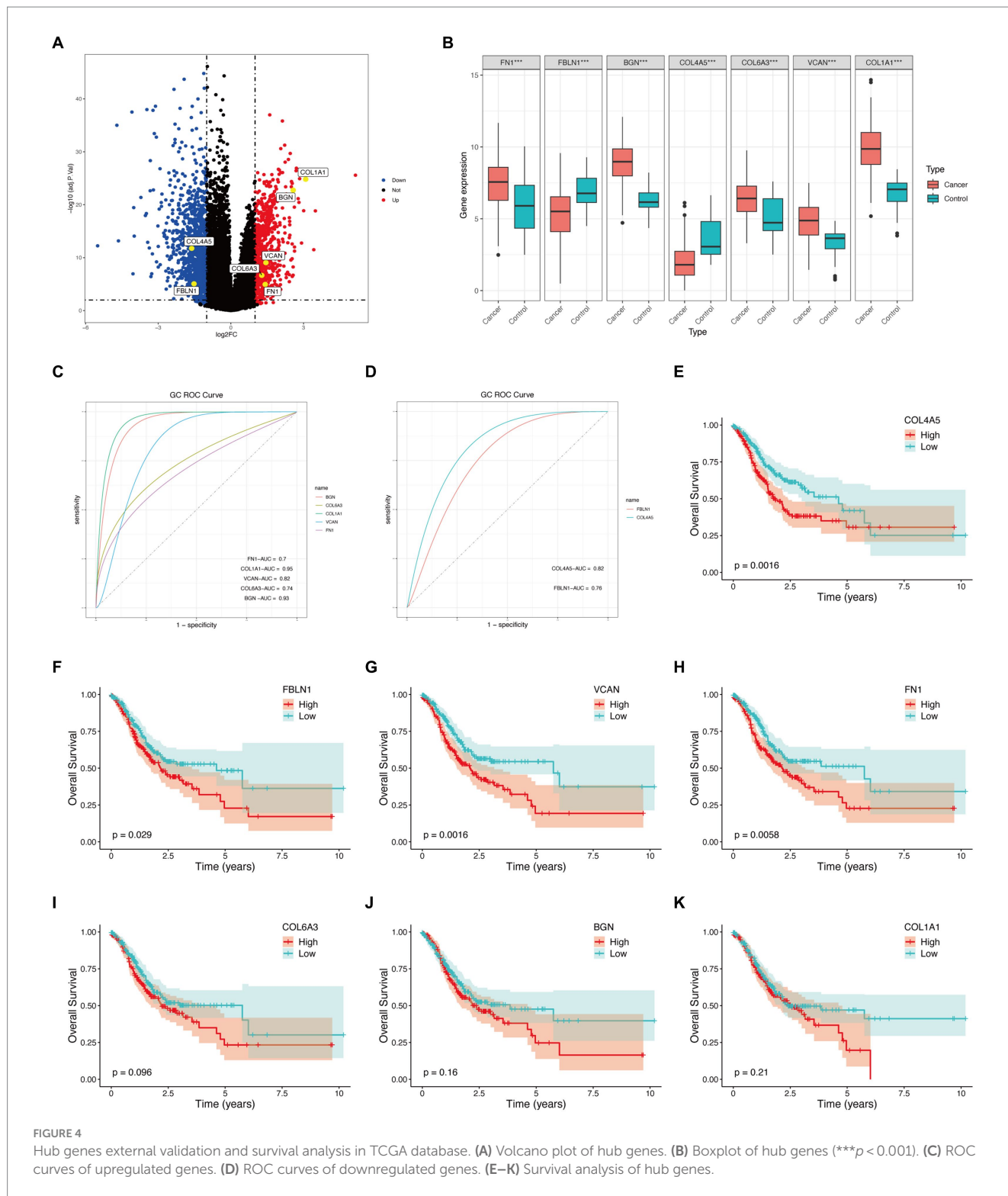
FIGURE 3
Functional enrichment analysis and PPI network of crosstalk genes. **(A)** The Venn diagram depicts 104 crosstalk genes formed by the intersection of representative genes from the two diseases. **(B)** KEGG analysis. **(C–E)** GO analysis. **(F)** Identification of candidate hub genes.

4 Discussion

A meta-analysis revealed that T2D could raise the incidence of GC by 19% (49). In addition, metformin, a common medicine used in the treatment of T2D (50), reduced the risk of GC by 24% (51).
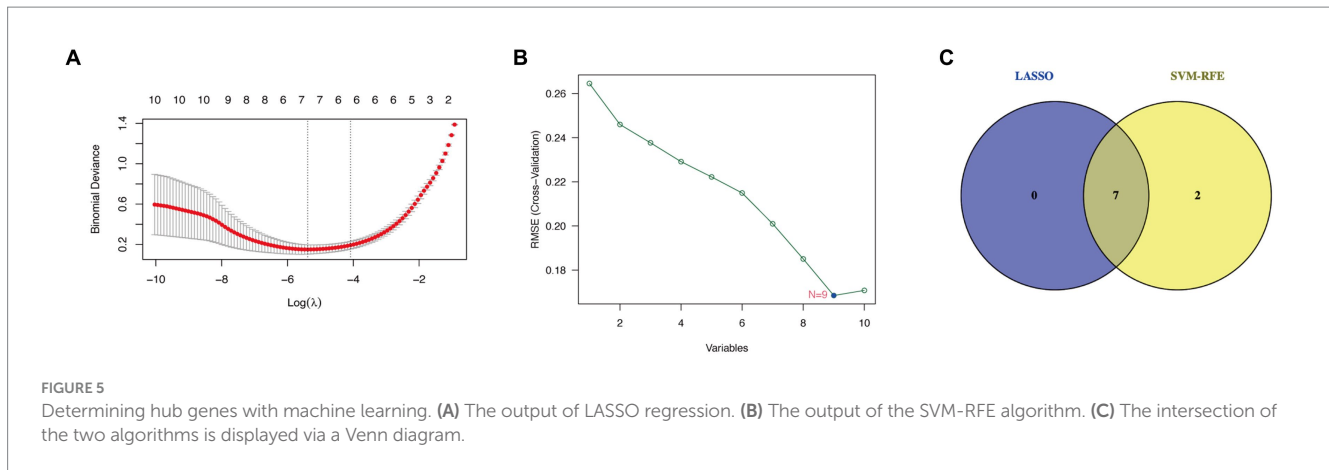
Meanwhile, a number of cancer types—particularly those of the pancreas, colon, breast, and stomach—can raise the chance of developing diabetes (7, 52–54). GC and T2D may be co-drivers of each other. The molecular processes driving the complicated interactions between these two diseases remain unexplained. This is the first study investigating the shared genes and common signatures

FIGURE 4

Hub genes external validation and survival analysis in TCGA database. **(A)** Volcano plot of hub genes. **(B)** Boxplot of hub genes (***$p < 0.001$). **(C)** ROC curves of upregulated genes. **(D)** ROC curves of downregulated genes. **(E–K)** Survival analysis of hub genes.

of GC and T2D through bioinformatic analysis and machine learning algorithms in an effort to facilitate early detection, improved treatment, and prompt prevention.

According to the GO analysis, we found the CC and MF of crosstalk genes between GC and T2D were mainly focused on extracellular matrix (ECM), while the KEGG analysis also verified crosstalk genes were enriched in "ECM-receptor interaction." ECM is

a crucial element of the GC microenvironment (55) as well as a major tumorigenesis regulator (56). Excessive deposition of ECM is one of the hallmarks of poor cancer prognosis (57). By stimulating cellular mechanisms associated with cell metabolic control, angiogenesis, and ECM receptors, gastric ECM remodeling promotes the growth of tumors (58, 59). Meanwhile, inappropriate deposition of ECM proteins is associated with multiple complications of diabetes, such as

FIGURE 5
Determining hub genes with machine learning. **(A)** The output of LASSO regression. **(B)** The output of the SVM-RFE algorithm. **(C)** The intersection of the two algorithms is displayed via a Venn diagram.

delayed wound healing (60), diabetic retinopathy (61, 62), diabetic foot ulcers (63), and diabetic nephropathy (DN) (64). In addition, hyperglycemia accompanying diabetes can lead to the impaired degradation or synthesis of HSPG in ECM cells, resulting in DN, cardiovascular disease, and retinopathy (65). Therefore, we hypothesized that ECM might be crucial to the comorbidity of T2D and GC.

"Aging" was in first place in the BP of the GO analysis. Cellular senescence is both a cause and a consequence of T2D (66, 67). Numerous senescence-associated secretory phenotypic factors released by senescent cells can mediate the dysfunction of pancreatic β-cell and adipose tissue, as well as insulin resistance, contributing to the pathogenesis of T2D (67, 68). In turn, hyperglycemia and metabolic changes can stimulate senescent cell formation (69), leading to various diabetic complications (70). In addition, senescence contributes to both the prevention and progression of tumors (71). Although senescence has been proven to be a conserved tumor suppressor mechanism (72, 73), the senescence-associated secretory phenotype can also facilitate tumor cell growth (74, 75).

"Cellular response to copper ion" ranked second in BP. Copper ion is associated with the development of T2D and may be a therapeutic target (76). In parallel with this, copper has garnered significant interest in the field of cancer treatment since it may be a limiting element in numerous facets of cancer advancement, such as angiogenesis, proliferation, and metastasis (77, 78). Hence, "cellular response to copper ion" could be the common BP of T2D and GC.

Through KEGG analysis, the "AGE-RAGE signaling pathway in diabetic complications" was considered a bidirectional pathway of T2D and GC. AGEs are synthesized at an accelerated rate during hyperglycemia, and RAGE is an advanced glycation end-product receptor (79). AGEs mainly trigger signaling pathways through RAGE that lead to cellular stress and dysfunction, and harm target organs, resulting in complications (80). Many diabetic complications are associated with AGE-RAGE signaling pathways, such as DN (81), cardiovascular disease (80, 82), and vascular calcification (79). Currently, it has been established that AGE-RAGE signaling contributes to the growth of different cancer (83). AGE-RAGE signaling may facilitate crosstalk between cancer microenvironment components and cells, inducing hypoxia, autophagy, endoplasmic reticulum stress, mitochondrial dysfunction, and epigenetic modification, suggesting that the AGE-RACE signal is an essential driving factor in cancer development (84). According to a study, the

AGE-RAGE signaling pathway may generate a positive feedback loop with oxidative stress, increasing the likelihood of cancer in people with diabetes (85).

Through multiple bioinformatics methods, seven hub genes were discovered to be implicated in the co-pathogenesis of GC and T2D. All the hub genes are connected with GC and may be poor prognostic markers in GC and many other cancers (86–92). BGN (93) and FN1 (94) may stimulate GC cell proliferation, invasion, migration, and EMT, which facilitate tumor progression. Aberrant VCAN expression is associated with modifications in ECM homeostasis, cell adhesion, differentiation, and proliferation (95, 96), thereby contributing to the carcinogenic potential of GC (97, 98). FBLN1 is identified as a candidate tumor suppressor gene whose inactivation can contribute to gastric carcinogenesis (89). COL4A5 is an independent prognostic marker for GC, especially in diffuse-type GC (90, 99). COL1A1 was found to be one of the cancer-associated fibroblast (CAF) markers for GC and a poor prognostic signature gene for CAF infiltration (100). COL1A1 is possibly a useful molecular marker and therapeutic target for GC (91). Overexpression of COL6A3 in GC promotes tumor growth and progression (92). It affects the tumor microenvironment, thereby promoting tumor inflammation and angiogenesis (101). As for T2D, FBLN1 is associated with mortality in T2D patients (102). COL1A1 is linked to hypoglycemic activity and is a novel therapeutic target that may be used to treat T2D (103). Although not all hub genes are well understood in T2D, they may be associated with diabetes because they are related to the ECM. VCAN is known to support ECM homeostasis (104). FN1 is an accumulation constituent of the ECM in the case of hyperglycemia (105). FN1 (106), COL1A1 (103) and COL4A5 (107) are significantly correlated with ECM–receptor interaction. COL4A5 (108) and BGN (109) are established to be contributors to the production of excess ECM, regulating ECM deposition. FBLN1 is an ECM protein (110, 111). COL6A3 encodes type VI collagen, which is found in the ECM of practically every tissue (112, 113). Moreover, the hub genes are connected with the development of diabetes complications, such as DN (96, 107, 114, 115), cardiovascular disease (116), and vascular stiffness (117).

In this study, ROC curves suggested that all seven hub genes had a good predictive effect on the occurrence of GC. However, survival analysis revealed that only four genes (VCAN, FN1, FBLN1, and COL4A5) were intimately connected to the GC prognosis. Around the sixth year, there was a reversal in the overall survival rate of the high-risk and low-risk COL4A5 groups. Among them, the overall survival of the

high-risk and low-risk groups of COL4A5 was reversed around the sixth year. Nevertheless, GC has a notoriously bad prognosis. According to statistics, less than 20% of patients with advanced cancer infections will survive for 5 years (118). Therefore, we considered that COL4A5 still had good predictive performance. However, further *in vitro* and *in vivo* studies are needed to validate our findings on the functional roles of the identified hub genes in GC. This validation will elucidate molecular mechanisms, validate prognostic significance, and guide the development of diagnostic and therapeutic approaches. Additionally, future research should focus on confirming the prognostic significance of COL4A5 and other genes in independent patient cohorts and exploring their potential as biomarkers for early detection and personalized treatment strategies. These efforts are crucial for translating our findings into clinically relevant applications to improve patient outcomes in GC management.

Despite our speculation on the potential mechanism of the connection between GC and T2D, certain limitations persist. First, only a few corresponding clinical data could be obtained from public databases. In addition, we failed to find patients with both GC and T2D for research. Finally, wet experiments could not be conducted to confirm our results because of the constraints of laboratory conditions.

5 Conclusion

This research revealed that "ECM-receptor interaction," "AGE-RAGE signaling pathway in diabetic complications," "aging," and "cellular response to copper ion" are the possible crosstalk mechanisms of GC and T2D. Additionally, it identified seven genes (BGN, VCAN, FN1, FBLN1, COL4A5, COL1A1, and COL6A3) as shared genes and potential targets for treatment in individuals with both GC and T2D. Notably, COL4A5 exhibited a reversal in overall survival rates around the sixth year. However, further investigation is warranted to confirm these conclusions.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

BX: Conceptualization, Data curation, Funding acquisition, Methodology, Resources, Software, Validation, Visualization, Writing – review & editing. PZ: Conceptualization, Data curation, Methodology, Resources, Software, Supervision, Validation, Visualization, Writing – review & editing. YX: Conceptualization, Data curation, Funding acquisition, Investigation, Resources, Software, Supervision, Validation, Visualization, Writing – review & editing. QL: Data curation, Formal

analysis, Investigation, Methodology, Project administration, Writing – review & editing. JXi: Data curation, Formal analysis, Methodology, Project administration, Software, Visualization, Writing – review & editing. JXu: Conceptualization, Formal analysis, Funding acquisition, Methodology, Project administration, Resources, Software, Supervision, Visualization, Writing – review & editing. WW: Conceptualization, Data curation, Formal analysis, Software, Supervision, Validation, Visualization, Writing – review & editing. XY: Conceptualization, Formal analysis, Funding acquisition, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmed.2024.1382004/full#supplementary-material

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* (2021) 71:209–49. doi: 10.3322/caac.21660

2. Wang Q, Zhang X, Zhong Y, Wei S, Li L, Wei W, et al. Clinical characteristics and changing trajectories of esophageal cancer and gastric cancer in China from 2010 to

2019: an analysis of a hospital-based database of 24,327 patients. *Front Oncol.* (2023) 13:1126841. doi: 10.3389/fonc.2023.1126841

3. Thrift AP, Wenker TN, El-Serag HB. Global burden of gastric cancer: epidemiological trends, risk factors, screening and prevention. *Nat Rev Clin Oncol.* (2023) 20:338–49. doi: 10.1038/s41571-023-00747-0

4. Hu Y, Zeng N, Ge Y, Wang D, Qin X, Zhang W, et al. Identification of the shared gene signatures and biological mechanism in type 2 diabetes and pancreatic cancer. *Front Endocrinol (Lausanne)*. (2022) 13:847760. doi: 10.3389/fendo.2022.847760

5. Sun H, Saeedi P, Karuranga S, Pinkepank M, Ogurtsova K, Duncan BB, et al. IDF diabetes atlas: global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045. *Diabetes Res Clin Pract*. (2022) 183:109119. doi: 10.1016/j.diabres.2021.109119

6. ALi A, Manzoor U, Ali SL, Marsool MD, Parida PK, Marsool AD, et al. Currently trending and futuristic biological modalities in the management of different types of diabetes: a comprehensive review. *J Popul Ther Clin Pharmacol*. (2023) 30:2948–70. doi: 10.53555/jptcp.v30i18.3467

7. Hwangbo Y, Kang D, Kang M, Kim S, Lee EK, Kim YA, et al. Incidence of diabetes after cancer development a korean national cohort study. *JAMA Oncol*. (2018) 4:1099–105. doi: 10.1001/jamaoncol.2018.1684

8. Tsilidis KK, Kasimis JC, Lopez DS, Ntzani EE, Ioannidis JP. Type 2 diabetes and cancer: umbrella review of meta-analyses of observational studies. *BMJ*. (2015) 350:g7607. doi: 10.1136/bmj.g7607

9. Currie CJ, Poole CD, Jenkins-Jones S, Gale EAM, Johnson JA, Morgan CL. Mortality after incident cancer in people with and without type 2 diabetes impact of metformin on survival. *Diabetes Care*. (2012) 35:299–304. doi: 10.2337/dc11-1313

10. Cheung KS, Chan EW, Chen L, Seto WK, Wong ICK, Leung WK. Diabetes increases risk of gastric cancer after *Helicobacter pylori* eradication: a territory-wide study with propensity score analysis. *Diabetes Care*. (2019) 42:1769–75. doi: 10.2337/dc19-0437

11. Kwon Y, Ha J, Kim D, Hwang J, Park SH, Kwon JW, et al. The association between weight change after gastric cancer surgery and type 2 diabetes risk: a nationwide cohort study. *J Cachexia Sarcopeni*. (2023) 14:826–34. doi: 10.1002/jcsm.13206

12. Tseng CH. Diabetes conveys a higher risk of gastric cancer mortality despite an age-standardised decreasing trend in the general population in Taiwan. *Gut*. (2011) 60:774–9. doi: 10.1136/gut.2010.226522

13. Zhou YF, Shi WX, Zhao D, Xiao SJ, Wang K, Wang J. Identification of immune-associated genes in diagnosing aortic valve calcification with metabolic syndrome by integrated bioinformatics analysis and machine learning. *Front Immunol*. (2022) 13:937886. doi: 10.3389/fimmu.2022.937886

14. Dinh P, Tran C, Dinh T, Ali A, Pan S. Hsa_CircRNA_0000284 acts as a ceRNA to participate in coronary heart disease progression by sponging miRNA-338-3p via regulating the expression of ETS_1. *J Biomol Struct Dyn*. (2023) 42:5114–27. doi: 10.1080/07391102.2023.2225109

15. Ali SL, Ali A, Alamri A, Baiduissenova A, Dusmagambetov M, Abduldayeva A. genomic annotation for vaccine target identification and immunoinformatics-guided multi-epitope-based vaccine design against songling virus through screening its whole genome encoded proteins. *Front Immunol*. (2023) 14:1284366. doi: 10.3389/fimmu.2023.1284366

16. Manzoor U, Ali A, Ali SL, Abdelkarem O, Kanwal S, Alotaibi SS, et al. Mutational screening of GDAP1 in dysphonia associated with charcot-marie-tooth disease: clinical insights and phenotypic effects. *J Genet Eng Biotechnol*. (2023) 21:119. doi: 10.1186/s43141-023-00568-9

17. Luo JF, Pan M, Mo K, Mao YW, Zou DH. Emerging role of artificial intelligence in diagnosis, classification and clinical management of glioma. *Semin Cancer Biol*. (2023) 91:110–23. doi: 10.1016/j.semcancer.2023.03.006

18. Fajarda O, Almeida JR, Duarte-Pereira S, Silva RM, Oliveira JL. Methodology to identify a gene expression signature by merging microarray datasets. *Comput Biol Med*. (2023) 159:106867. doi: 10.1016/j.compbiomed.2023.106867

19. Shafiq N, Arshad M, Ali A, Rida F, Mohany M, Arshad U, et al. Integrated computational modeling and *in-silico* validation of flavonoids-alliuocide G and alliuocide a as therapeutic agents for their multi-target potential: combination of molecular docking, MM-GBSA, ADMET and DFT analysis. *S Afr J Bot*. (2024) 169:276–300. doi: 10.1016/j.sajb.2024.04.031

20. 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*. (2015) 43:e47. doi: 10.1093/nar/gkv007

21. Robinson MD, McCarthy DJ, Smyth GK. Edger: A bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. (2010) 26:139–40. doi: 10.1093/bioinformatics/btp616

22. Wickham H. *Ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag (2016).

23. Kolde R. *Pheatmap: Pretty Heatmaps*. (2019). Available at: https://CRAN.R-project.org/package=pheatmap

24. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*. (2008) 9:559. doi: 10.1186/1471-2105-9-559

25. Zhang B, Horvath S. A general framework for weighted gene co-expression network analysis. *Stat Appl Genet Mol Biol*. (2005) 4:Article 17. doi: 10.2202/1544-6115.1128

26. Carbon S, Douglass E, Dunn N, Good B, Harris NL, Lewis SE, et al. The gene ontology resource: 20 years and still going strong. *Nucleic Acids Res*. (2019) 47:D330–8. doi: 10.1093/nar/gky1055

27. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*. (2000) 28:27–30. doi: 10.1093/nar/28.1.27

28. Wickham H. *Stringr: Simple, Consistent Wrappers for Common String Operations*. (2022). Available at: https://CRAN.R-project.org/package=stringr

29. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. String V11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res*. (2019) 47:D607–13. doi: 10.1093/nar/gky1131

30. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. (2003) 13:2498–504. doi: 10.1101/gr.1239303

31. Kumar SU, Kumar DT, Bithia R, Sankar S, Magesh R, Sidenna M, et al. Analysis of differentially expressed genes and molecular pathways in familial hypercholesterolemia involved in atherosclerosis: a systematic and bioinformatics approach. *Front Genet*. (2020) 11:734. doi: 10.3389/fgene.2020.00734

32. Balasundaram A, Udhaya Kumar S, George Priya Doss C. A computational model revealing the immune-related hub genes and key pathways involved in rheumatoid arthritis (RA). *Adv Protein Chem Struct Biol*. (2022) 129:247–73. doi: 10.1016/bs.apcsb.2021.11.006

33. Udhaya Kumar S, Balasundaram A, Anu Preethi V, Chatterjee S, Kameshwari Gollakota GV, Kashyap MK, et al. Integrative ontology and pathway-based approach identifies distinct molecular signatures in transcriptomes of esophageal squamous cell carcinoma. *Adv Protein Chem Struct Biol*. (2022) 131:177–206. doi: 10.1016/bs.apcsb.2022.04.003

34. Udhaya Kumar S, Datta A, Gnanasambandan R, Younes S, Medha T, Siva R, et al. Investigation of differentially expressed genes and dysregulated pathways involved in multiple sclerosis. *Adv Protein Chem Struct Biol*. (2022) 131:235–59. doi: 10.1016/bs.apcsb.2022.05.003

35. Sanz H, Valim C, Vegas E, Oller JM, Reverter F. SVM-RFE: selection and visualization of the most relevant features through non-linear kernels. *BMC Bioinformatics*. (2018) 19:432. doi: 10.1186/s12859-018-2451-4

36. Zhao ZY, He SJ, Yu XC, Lai XF, Tang S, Mariya ME, et al. Analysis and experimental validation of rheumatoid arthritis innate immunity gene CYFIP2 and pan-cancer. *Front Immunol*. (2022) 13:954848. doi: 10.3389/fimmu.2022.954848

37. Karatzoglou A, Smola A, Hornik K, Zeileis A. Kernlab - an S4 package for kernel methods in R. *J Stat Softw*. (2004) 11:1–20. doi: 10.18637/jss.v011.i09

38. Karatzoglou A, Smola A, Hornik K. *Kernlab: Kernel-Based Machine Learning Lab*. (2023). Available at: https://cran.r-project.org/web//packages/kernlab/kernlab.pdf

39. Kuhn M. Building predictive models in R using the caret package. *J Stat Softw*. (2008) 28:1–26. doi: 10.18637/jss.v028.i05

40. Meyer D, Dimitriadou E, Hornik K, Weingessel A, Leisch F. *E1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), Tu Wien*. (2023). Available at: https://CRAN.R-project.org/package=e1071

41. Yang C, Delcher C, Shenkman E, Ranka S. Machine learning approaches for predicting high cost high need patient expenditures in health care. *Biomed Eng Online*. (2018) 17:131. doi: 10.1186/s12938-018-0568-3

42. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw*. (2010) 33:1–22. doi: 10.18637/jss.v033.i01

43. Wickham H. Reshaping data with the reshape package. *J Stat Softw*. (2007) 21:1–20. doi: 10.18637/jss.v021.i12

44. Kassambara A. *Ggpubr: 'Ggplot2' Based Publication Ready Plots*. (2023). Available at: https://CRAN.R-project.org/package=ggpubr

45. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, et al. PROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*. (2011) 12:77. doi: 10.1186/1471-2105-12-77

46. Therneau T. *A Package for Survival Analysis in R*. (2023). Available at: https://CRAN.R-project.org/package=survival

47. Therneau TM, Grambsch PM. *Modeling Survival Data: Extending the Cox Model*. New York: Springer-Verlag (2000).

48. Kassambara A, Kosinski M, Biecek P. *Survminer: Drawing Survival Curves Using 'Ggplot2'*. (2021). Available at: https://CRAN.R-project.org/package=survminer

49. Yoon JM, Son KY, Eom CS, Durrance D, Park SM. Pre-existing diabetes mellitus increases the risk of gastric cancer: a meta-analysis. *World J Gastroenterol*. (2013) 19:936–45. doi: 10.3748/wjg.v19.i6.936

50. Ferrannini E. The target of metformin in type 2 diabetes. *N Engl J Med*. (2014) 371:1547–8. doi: 10.1056/NEJMcibr1409796

51. Zhou XL, Xue WH, Ding XF, Li LF, Dou MM, Zhang WJ, et al. Association between metformin and the risk of gastric cancer in patients with type 2 diabetes mellitus: a meta-analysis of cohort studies. *Oncotarget*. (2017) 8:55622–31. doi: 10.18632/oncotarget.16973

52. Lipscombe LL, Chan WW, Yun L, Austin PC, Anderson GM, Rochon PA. Incidence of diabetes among postmenopausal breast cancer survivors. *Diabetologia*. (2013) 56:476–83. doi: 10.1007/s00125-012-2793-9

53. De Bruijn KMJ, van Eijck CHJ. New-onset diabetes after distal pancreatectomy a systematic review. *Ann Surg*. (2015) 261:854–61. doi: 10.1097/Sla.0000000000000819

54. Singh S, Earle CC, Bae SJ, Fischer HD, Yun LS, Austin PC, et al. Incidence of diabetes in colorectal cancer survivors. *J Natl Cancer Inst*. (2016) 108:djv402. doi: 10.1093/jnci/djv402

55. Quail DF, Joyce JA. Microenvironmental regulation of tumor progression and metastasis. *Nat Med*. (2013) 19:1423–37. doi: 10.1038/nm.3394

56. Bonnans C, Chou J, Werb Z. Remodelling the extracellular matrix in development and disease. *Nat Rev Mol Cell Biol*. (2014) 15:786–801. doi: 10.1038/nrm3904

57. Levental KR, Yu HM, Kass L, Lakins JN, Egeblad M, Erler JT, et al. Matrix crosslinking forces tumor progression by enhancing integrin signaling. *Cell*. (2009) 139:891–906. doi: 10.1016/j.cell.2009.10.027

58. Moreira AM, Ferreira RM, Carneiro P, Figueiredo J, Osorio H, Barbosa J, et al. Proteomic identification of a gastric tumor ecm signature associated with cancer progression. *Front Mol Biosci*. (2022) 9:818552. doi: 10.3389/fmolb.2022.818552

59. Nwanna E, Ojo R, Shafiq N, Ali A, Okello E, Oboh G. An *in silico in vitro* and *in vivo* study on the influence of an eggplant fruit (*Solanum Anguivi* Lam) diet on metabolic dysfunction in the sucrose-induced diabetic-like fruit fly (*Drosophila Melanogaster*). *Food Secur*. (2024) 13:559. doi: 10.3390/foods13040559

60. Liu Y, Liu YQ, He WJ, Mu XR, Wu XQ, Deng JY, et al. Fibroblasts: immunomodulatory factors in refractory diabetic wound healing. *Front Immunol*. (2022) 13:918223. doi: 10.3389/fimmu.2022.918223

61. Uemura A, Kusuhara S, Wiegand SJ, Yu RT, Nishikawa SI. TLX acts as a proangiogenic switch by regulating extracellular assembly of fibronectin matrices in retinal astrocytes. *J Clin Invest*. (2006) 116:369–77. doi: 10.1172/Jci25964

62. Ai J, Liu Y, Sun JH. Advanced glycation end-products stimulate basic fibroblast growth factor expression in cultured muller cells. *Mol Med Rep*. (2013) 7:16–20. doi: 10.3892/mmr.2012.1152

63. Maione AG, Smith A, Kashpur O, Yanez V, Knight E, Mooney DJ, et al. Altered ECM deposition by diabetic foot ulcer-derived fibroblasts implicates fibronectin in chronic wound repair. *Wound Repair Regen*. (2016) 24:630–43. doi: 10.1111/wrr.12437

64. Liu LL, Wang YY, Yan R, Liang LQ, Zhou XC, Liu HM, et al. BMP-7 inhibits renal fibrosis in diabetic nephropathy via MIR-21 downregulation. *Life Sci*. (2019) 238:238. doi: 10.1016/j.lfs.2019.116957

65. Hiebert LM. Heparan sulfate proteoglycans in diabetes. *Semin Thromb Hemost*. (2021) 47:261–73. doi: 10.1055/s-0041-1724118

66. Narasimhan A, Flores RR, Robbins PD, Niedernhofer LJ. Role of cellular senescence in type II diabetes. *Endocrinology*. (2021) 162:162. doi: 10.1210/endocr/bqab136

67. Palmer AK, Tchkonia T, LeBrasseur NK, Chini EN, Xu M, Kirkland JL. Cellular Senescence in type 2 diabetes: a therapeutic opportunity. *Diabetes*. (2015) 64:2289–98. doi: 10.2337/db14-1820

68. Tchkonia T, Morbeck DE, Von Zglinicki T, Van Deursen J, Lustgarten J, Scrable H, et al. Fat tissue, aging, and cellular senescence. *Aging Cell*. (2010) 9:667–84. doi: 10.1111/j.1474-9726.2010.00608.x

69. Regulski M. Understanding diabetic induction of cellular senescence: a concise review. *Wounds*. (2018) 30:96–101.

70. Palmer AK, Gustafson B, Kirkland JL, Smith U. Cellular senescence: at the nexus between ageing and diabetes. *Diabetologia*. (2019) 62:1835–41. doi: 10.1007/s00125-019-4934-x

71. Lasry A, Ben-Neriah Y. Senescence-associated inflammatory responses: aging and cancer perspectives. *Trends Immunol*. (2015) 36:217–28. doi: 10.1016/j.it.2015.02.009

72. van Deursen JM. The role of senescent cells in ageing. *Nature*. (2014) 509:439–46. doi: 10.1038/nature13193

73. Di Micco R, Krizhanovsky V, Baker D, di Fagagna FD. Cellular senescence in ageing: from mechanisms to therapeutic opportunities. *Nat Rev Mol Cell Biol*. (2021) 22:75–95. doi: 10.1038/s41580-020-00314-w

74. Pribluda A, Elyada E, Wiener Z, Hamza H, Goldstein RE, Biton M, et al. A senescence-inflammatory switch from cancer-inhibitory to cancer-promoting mechanism. *Cancer Cell*. (2013) 24:242–56. doi: 10.1016/j.ccr.2013.06.005

75. Coppe JP, Desprez PY, Krtolica A, Campisi J. The senescence-associated secretory phenotype: the dark side of tumor suppression. *Annu Rev Pathol Mech*. (2010) 5:99–118. doi: 10.1146/annurev-pathol-121808-102144

76. Tanaka A, Kaneto H, Miyatsuka T, Yamamoto K, Yoshiuchi K, Yamasaki Y, et al. Role of copper ion in the pathogenesis of type 2 diabetes. *Endocr J*. (2009) 56:699–706. doi: 10.1507/endocrj.K09E-051

77. Denoyer D, Clatworthy SAS, Cater MA. Copper complexes in cancer therapy. *Met ions. Life Sci*. (2018) 18:18. doi: 10.1515/9783110470734-022

78. da Silva DA, De Luca A, Squitti R, Rongioletti M, Rossi L, Machado CML, et al. Copper in tumors and the use of copper-based compounds in cancer treatment. *J Inorg Biochem*. (2022) 226:111634. doi: 10.1016/j.jinorgbio.2021.111634

79. Kay AM, Simpson CL, Stewart JA Jr. The role of age/rage signaling in diabetes-mediated vascular calcification. *J Diabetes Res*. (2016) 2016:6809703. doi: 10.1155/2016/6809703

80. Ramasamy R, Yan SF, Schmidt AM. Receptor for age (rage): signaling mechanisms in the pathogenesis of diabetes and its complications. *Ann N Y Acad Sci*. (2011) 1243:88–102. doi: 10.1111/j.1749-6632.2011.06320.x

81. Pathomthongtaweechai N, Chutipongtanate S. Age/rage signaling-mediated endoplasmic reticulum stress and future prospects in non-coding rna therapeutics for diabetic nephropathy. *Biomed Pharmacother*. (2020) 131:110655. doi: 10.1016/j.biopha.2020.110655

82. Burr SD, Harmon MB, Jr JAS. The impact of diabetic conditions and age/rage signaling on cardiac fibroblast migration. *Front Cell Dev Biol*. (2020) 8:112. doi: 10.3389/fcell.2020.00112

83. Waghela BN, Vaidya FU, Ranjan K, Chhipa AS, Tiwari BS, Pathak C. Age-rage synergy influences programmed cell death signaling to promote cancer. *Mol Cell Biochem*. (2021) 476:585–98. doi: 10.1007/s11010-020-03928-y

84. Muthyalaiah YS, Jonnalagadda B, John CM, Arockiasamy S. Impact of advanced glycation end products (ages) and its receptor (rage) on cancer metabolic signaling pathways and its progression. *Glycoconj J*. (2021) 38:717–34. doi: 10.1007/s10719-021-10031-x

85. Abe R, Yamagishi S. Age-rage system and carcinogenesis. *Curr Pharm Des*. (2008) 14:940–5. doi: 10.2174/138161208784139765

86. Zhao SF, Yin XJ, Zhao WJ, Liu LC, Wang ZP. Biglycan as a potential diagnostic and prognostic biomarker in multiple human cancers. *Oncol Lett*. (2020) 19:1673–82. doi: 10.3892/ol.2020.11266

87. Li WF, Han F, Fu M, Wang ZQ. High expression of vcan is an independent predictor of poor prognosis in gastric cancer. *J Int Med Res*. (2020) 48:48. doi: 10.1177/0300060519891271

88. Wang H, Zhang JC, Li H, Yu H, Chen SY, Liu SH, et al. FN1 is a prognostic biomarker and correlated with immune infiltrates in gastric cancers. *Front Oncol*. (2022) 12:918719. doi: 10.3389/fonc.2022.918719

89. Cheng YY, Jin H, Liu X, Siu JMT, Wong YP, Ng EKO, et al. Fibulin 1 is downregulated through promoter hypermethylation in gastric cancer. *Brit J Cancer*. (2008) 99:2083–7. doi: 10.1038/sj.bjc.6604760

90. Zeng X, Wang HY, Wang YP, Bai SY, Pu K, Zheng Y, et al. COL4A family: potential prognostic biomarkers and therapeutic targets for gastric cancer. *Transl Cancer Res*. (2020) 9:5218–32. doi: 10.21037/tcr-20-517

91. Li YM, Luo Y, Tian Q, Lai Y, Xu L, Yun HL, et al. Integrated bioinformatics analysis for identifying the significant genes as poor prognostic markers in gastric adenocarcinoma. *J Oncol*. (2022) 2022:1–12. doi: 10.1155/2022/9080460

92. Xie X, Liu X, Zhang Q, Yu J. Overexpression of collagen VI α3 in gastric cancer. *Oncol Lett*. (2014) 7:1537–43. doi: 10.3892/ol.2014.1910

93. Wu HT, Xiang ZX, Huang GQ, He QM, Song JL, Dou RZ, et al. BGN/FAP/STAT3 positive feedback loop mediated mutual interaction between tumor cells and mesothelial cells contributes to peritoneal metastasis of gastric cancer. *Int J Biol Sci*. (2023) 19:465–83. doi: 10.7150/ijbs.72218

94. Zhang HC, Sun ZG, Li Y, Fan D, Jiang H. MicroRNA-200c binding to FN1 suppresses the proliferation, migration and invasion of gastric cancer cells. *Biomed Pharmacother*. (2017) 88:285–92. doi: 10.1016/j.biopha.2017.01.023

95. Wight TN. Versican: a versatile extracellular matrix proteoglycan in cell biology. *Curr Opin Cell Biol*. (2002) 14:617–23. doi: 10.1016/S0955-0674(02)00375-7

96. Xu QN, Li BJ, Wang YC, Wang CL, Feng S, Xue L, et al. Identification of VCAN as hub gene for diabetic kidney disease immune injury using integrated bioinformatics analysis. *Front Physiol*. (2021) 12:651690. doi: 10.3389/fphys.2021.651690

97. Jiang KY, Liu HM, Xie DY, Xiao Q. differentially expressed genes ASPN, COL1A1, FN1, VCAN AND MUC5AC are potential prognostic biomarkers for gastric cancer. *Oncol Lett*. (2019) 17:3191–202. doi: 10.3892/ol.2019.9952

98. Zhang ZW, Zhang JP, Miao L, Liu K, Yang SS, Pan CY, et al. Interleukin-11 promotes the progress of gastric carcinoma via abnormally expressed versican. *Int J Biol Sci*. (2012) 8:383–93. doi: 10.7150/ijbs.3579

99. Bao BW, Zheng CL, Yang BW, Jin Y, Hou KZ, Li Z, et al. Identification of subtype-specific three-gene signature for prognostic prediction in diffuse type gastric cancer. *Front Oncol*. (2019) 9:1243. doi: 10.3389/fonc.2019.01243

100. Metin CU, Ozcan G. Comprehensive bioinformatic analysis reveals a cancer-associated fibroblast gene signature as a poor prognostic factor and potential therapeutic target in gastric cancer. *BMC Cancer*. (2022) 22:692. doi: 10.1186/s12885-022-09736-5

101. Chen P, Cescon M, Bonaldo P. Collagen VI in cancer and its biological mechanisms. *Trends Mol Med*. (2013) 19:410–7. doi: 10.1016/j.molmed.2013.04.001

102. Laugesen E, Hoyem P, Christiansen JS, Knudsen ST, Hansen KW, Argraves WS, et al. Plasma levels of the arterial wall protein fibulin-1 are associated with carotid-femoral pulse wave velocity: a cross-sectional study. *Cardiovasc Diabetol*. (2013) 12:107. doi: 10.1186/1475-2840-12-107

103. Lin GP, Wan XZ, Liu D, Wen YX, Yang CF, Zhao C. COL1A1 as a potential new biomarker and therapeutic target for type 2 diabetes. *Pharmacol Res*. (2021) 165:165. doi: 10.1016/j.phrs.2021.105436

104. Ricciardelli C, Sakko A, Ween M, Russell D, Horsfall D. The biological role and regulation of versican levels in cancer. *Cancer Metast Rev*. (2009) 28:233–45. doi: 10.1007/s10555-009-9182-y

105. Wang YJ, Zhao MM, Zhang Y. Identification of fibronectin 1 (FN1) and complement component 3 (C3) as immune infiltration-related biomarkers for diabetic nephropathy using integrated bioinformatic analysis. *Bioengineered*. (2021) 12:5386–401. doi: 10.1080/21655979.2021.1960766

106. Li ZK, Feng JX, Zhong JT, Lu MZ, Gao XJ, Zhang YF. Screening of the key genes and signalling pathways for diabetic nephropathy using bioinformatics analysis. *Front Endocrinol*. (2022) 13:864407. doi: 10.3389/fendo.2022.864407

107. Yang FY, Cui ZH, Deng HJ, Wang Y, Chen Y, Li HQ, et al. Identification of miRNAs-genes regulatory network in diabetic nephropathy based on bioinformatics analysis. *Medicine*. (2019) 98:e16225. doi: 10.1097/MD.0000000000016225

108. Li YN, Zhong Y, Gong WJ, Gao XH, Qi HL, Liu K, et al. C-peptide prevents SMAD3 binding to alpha promoters to inhibit collagen type IV synthesis. *J Mol Endocrinol*. (2018) 61:47–56. doi: 10.1530/Jme-18-0009

109. Yu ML, He XJ, Song X, Gao JG, Pan JQ, Zhou TY, et al. Biglycan promotes hepatic fibrosis through activating heat shock protein 47. *Liver Int*. (2023) 43:500–12. doi: 10.1111/liv.15477

110. Godyna S, Mann DM, Argraves WS. A quantitative-analysis of the incorporation of fibulin-1 into extracellular-matrix indicates that fibronectin assembly is required. *Matrix Biol*. (1995) 14:467–77. doi: 10.1016/0945-053x(95)90004-7

111. Harikrishnan K, Cooley MA, Sugi Y, Barth JL, Rasmussen LM, Kern CB, et al. Fibulin-1 suppresses endothelial to mesenchymal transition in the proximal outflow tract. *Mech Develop*. (2015) 136:123–32. doi: 10.1016/j.mod.2014.12.005

112. Cescon M, Gattazzo F, Chen PW, Bonaldo P. Collagen VI at a glance. *J Cell Sci*. (2015) 128:3525–31. doi: 10.1242/jcs.169748

113. Lamande SR, Bateman JF. Collagen VI disorders: insights on form and function in the extracellular matrix and beyond. *Matrix Biol*. (2018) 71-72:348–67. doi: 10.1016/j.matbio.2017.12.008

114. Gao Z, S A, Li XM, Li XL, Sui LN. Identification of key candidate genes and chemical perturbagens in diabetic kidney disease using integrated bioinformatics analysis. *Front Endocrinol (Lausanne)*. (2021) 12:721202. doi: 10.3389/fendo.2021.721202

115. Tsai YC, Hung WW, Chang WA, Wu PH, Wu LY, Lee SC, et al. Autocrine exosomal Fibulin-1 as a target of MiR-1269b induces epithelial-mesenchymal transition in proximal tubule in diabetic nephropathy. *Front Cell Dev Biol*. (2021) 9:789716. doi: 10.3389/fcell.2021.789716

116. Scholze A, Bladbjerg EM, Sidelmann JJ, Diederichsen AC, Mickley H, Nybo M, et al. Plasma concentrations of extracellular matrix protein fibulin-1 are related to cardiovascular risk markers in chronic kidney disease and diabetes. *Cardiovasc Diabetol*. (2013) 12:6. doi: 10.1186/1475-2840-12-6

117. Paapstel K, Zilmer M, Eha J, Tootsi K, Piir A, Kals J. Association between fibulin-1 and aortic augmentation index in male patients with peripheral arterial disease. *Eur J Vasc Endovasc Surg*. (2016) 51:76–82. doi: 10.1016/j.ejvs.2015.09.004

118. Nagini S. Carcinoma of the stomach: a review of epidemiology, pathogenesis, molecular genetics and chemoprevention. *World J Gastrointest Oncol*. (2012) 4:156–69. doi: 10.4251/wjgo.v4.i7.156