RESEARCH

Integrative genomics unveils basement membrane-related diagnostic markers and therapeutic targets in esophageal squamous cell carcinoma

Han Zhang[1*†], Xia Zhang[1†], Zhenyu Huang[1] and Hao Zhang[2*]

Abstract

Background Esophageal squamous cell carcinoma (ESCC) is often diagnosed at advanced stages due to the inherent limitations of current screening methodologies. Central to evaluating tumor invasion and prognostic assessment in ESCC is the integrity of the basement membrane (BM). However, current research on the implications of BM-related genes (BMRGs) in diagnosing ESCC remains sparse.

Methods We performed a comprehensive analysis using single-cell RNA-sequencing (scRNA-seq) data from the Gene Expression Omnibus (GEO) database, alongside gene expression profiles acquired from GEO and The Cancer Genome Atlas (TCGA) databases. This identified differentially expressed BMRGs in ESCC. Employing LASSO, RF, and SVM-RFE, we selected potential BM biomarkers and crafted a diagnostic nomogram for ESCC, validated by ROC curves and AUC values. We also explored immune infiltration and biological mechanisms through consensus clustering and GSVA, and utilized single cell trajectory analysis and GSCALite to study gene distributions and pathways. In vitro experiments further elucidated the role of these genes in ESCC carcinogenesis.

Results Here, we discovered that ESCC cell types exhibited markedly elevated BM-related scores. Our analysis pinpointed seven BM genes upregulated and linked to immune infiltration, showcasing unique gene expression profiles and varying immune cell densities across the BM-related subtypes. Furthermore, a robust positive correlation was observed between these genes expression and EMT activity. The knockdown of BGN significantly suppressed cell proliferation, migration, invasion, while also augmenting cell viability following chemotherapy drug treatment.

Conclusion Our study identified seven key BMRGs (*BGN*, *LAMB3*, *SPARC*, *MMP1*, *LUM*, *COL4A1*, and *NELL2*) and established a diagnostic nomogram for ESCC. Of noteworthy significance is the discovery of BGN as a promising drug target, indicating a novel strategy for future clinical combination therapies in ESCC.

Keywords Esophageal squamous cell carcinoma, Basement membrane, Diagnostic biomarkers, Machine learning, *BGN*

†Han Zhang and Xia Zhang have equally contributed to this work.

*Correspondence:
Han Zhang
zhanghantz@163.com
Hao Zhang
hao2012590@163.com

Zhang *et al. Biology Direct*   (2024) 19:79

Page 2 of 20

Introduction

Esophageal cancer is a prevalent malignant carcinoma, ranking seventh in global incidence and sixth in cancer-related mortality [1]. It comprises two primary subtypes, esophageal adenocarcinoma (EAC) and esophageal squamous cell carcinoma (ESCC), distinguished by etiology and epidemiological distributions. With ESCC accounting for over 90%, China has the highest incidence rate of esophageal cancer globally [2, 3]. Unfortunately, ESCC often presents without symptoms in its early stages, leading to late diagnoses and a loss of opportunities for conventional treatment [4]. Hence, there is an urgent need to identify novel and reliable biomarkers to enhance early ESCC diagnosis and screening.

Invasion and metastasis are the leading causes of cancer-related morbidity and mortality [5, 6], constituting fundamental 'hallmarks of cancer' as described by Hanahan and Weinberg [7, 8]. Basement membranes (BMs) are intricate layers of extracellular matrix (ECM) that act as a barrier between tumor cells and the stroma, playing a crucial role in preventing malignant cell invasion and metastasis [9]. The ECM is a vital non-cellular component of the tumor microenvironment (TME), with the capacity to influence immune cell recruitment and the progression of tumors. The breaching of the epithelial BM by tumor cells during the metastatic phase is a critical point in their dissemination to distant organs [10]. The alteration of ECM architecture by neoplastic invasion is crucial to tumorigenesis, malignant cell metastasis, and the remodeling of the TME [11]. BMs are primarily composed of laminin, which facilitates cellular interaction and collagen IV, which maintains membrane structure [12, 13]. They also harbor growth factors, heparan sulfate proteoglycans, and nidogen [14]. The complex composition of BMs is essential for sustaining cell polarity, promoting cell adhesion, and enabling cell migration. The distinct protein expressions and densities of BMs in normal versus pathophysiological conditions, such as cancer, are particularly noteworthy. Furthermore, the upregulation of matrix metalloproteinases (MMPs), enzymes that can degrade BMs, is closely linked to tumor invasion and metastasis [15]. Extensive studies have uncovered significant correlations between BM-related genes (BMRGs) and a variety of malignancies, including breast cancer [16], lung adenocarcinomas [17], bladder cancer [18] and renal cell carcinoma [19].

ESCC is a multifaceted disease characterized by a complex, multi-step progression. It initiates in normal epithelial cells, evolves through basal cell hyperplasia, and escalates through multiple phases of intraepithelial neoplasia, culminating in an aggressive, invasive cancer [20]. This progression encompasses the epithelial-mesenchymal transition (EMT) within the basement membrane, facilitating metastasis through the lymphovascular system [21]. Spatial transcriptomic analyses on multistage ESCC samples have elucidated that aberrant interactions among epithelial cells in the basal layer initiate EMT and expedite [22]. In the realms of diagnosis and treatment, comprehending the role of BMs is indispensable. Carcinoma in situ, confined to the epithelium, can be managed effectively. However, once the disease breaches the BMs, it infiltrates deeper tissue. The relevance of BMs is further underscored in therapeutic approaches such as radiotherapy, chemotherapy, and immunotherapy. Consequently, identifying genes associated with BMs is paramount for precise diagnosis and treatment, averting the dire metastatic outcomes in ESCC.

In the current report, we commenced by procuring ESCC single-cell RNA-sequencing (scRNA-seq) data from the Gene Expression Omnibus (GEO) database, employing a set of hallmark genes to compute the BM scores. We then performed differential expression analysis of BM-related genes, contrasting ESCC with normal tissues, utilizing gene expression profiles from GEO and TCGA databases. Subsequently, we harnessed three distinct machine learning algorithms to pinpoint the seven potential biomarkers, around which we constructed a diagnostic nomogram model for the prediction of ESCC. The expression profiles and diagnostic accuracy of these seven hub genes were corroborated using external databases. Building upon preliminary evaluations that included assessments of immune cell infiltration, immune expression signatures and their correlation with EMT, we advanced to functionally delineate the role of the BGN gene in ESCC carcinogenesis through a series of experimental approaches.

Our study unveils novel diagnostic biomarkers for ESCC by identifying previously unreported BMRGs. The limitations of bulk technologies in discerning signals from heterogeneous cell populations are well recognized. In contrast, scRNA-seq offers unprecedented resolution in mapping cellular compositions within complex microenvironments. Leveraging this technology, we adopted an interdisciplinary strategy to identify candidate genes with enhanced diagnostic efficacy and specificity, surpassing the reliability of biomarkers identified in prior studies that did not undergo ROC analysis for diagnostic efficiency assessment [23, 24]. Furthermore, we have pinpointed a novel therapeutic target for ESCC through predictive drug sensitivity and in vitro cellular assays, laying the groundwork for personalized therapeutic strategies in ESCC.

Zhang *et al. Biology Direct*   (2024) 19:79

Page 3 of 20

Materials and methods
Data collection
We acquired four microarray datasets (GSE53625, GSE44021, GSE23400, GSE20347) and one scRNA-seq data (GSE188900) from the GEO database. Additionally, we obtained RNA expression profiles of ESCC samples (n = 81) from the TCGA database.

Analysis of scRNA-seq data and calculation of BM-related gene module score
The R package "Seurat" was used to filter, normalize and cluster the scRNA-seq raw data. We employed the universal manifold approximation and projection (UMAP) method for dimensionality reduction. To annotate cell clusters, we referred to cell markers obtained from the CellMarker and PanglaoDB databases. The AUCell and ssGSEA algorithms were utilized to calculate the BM scores. A total of 222 BM-related genes have been reported elsewhere, including genes encoding proteins of the BM matrix and components of the BM zone (Supplementary Table S1) [25]. However, in these scRNA-seq data, 10 of these genes were not detected. Subsequently, the BM-related gene module score was computed for each cell type using 212 hallmark genes. Additionally, Monocle was employed for trajectory analysis to predict cell differentiation and visualize gene expression profiles within each cell state.

Screening of BM-related biomarkers by machine learning
The R package "limma" was used to identify differentially expressed genes (DEGs) between the tumor and normal groups, applying thresholds of an adjusted $P < 0.05$ and | log2 FC |> 1. Initially, we screened for the intersection of BM-related DEGs in TCGA, GSE53625, and GSE44021. Then, three machine learning algorithms, namely the least absolute shrinkage and selection operator (LASSO) algorithm [26], support vector machine-recursive feature elimination (SVM-RFE) algorithm [27], and random forest (RF) algorithm [28], were employed to identify characteristic BMRGs. LASSO analysis was conducted using ten-fold cross-validated penalty parameters via the R package "glmnet" with the minimal lambda value considered optimal. The SVM-RFE classifier was implemented with the minimum cross-validation error using the "e1071" package. RF algorithm analysis was executed via the "randomForest" package and genes with a Mean Decrease Gini value greater than two designated as characteristic genes. Finally, receiver operating characteristic (ROC) analysis was performed using the "pROC" R package and the area under the curve (AUC) was used to estimate the diagnostic efficacy of candidate BM genes in both training and validation GEO datasets.

Functional enrichment analysis
The R package "clusterProfiler" was employed to perform Gene Ontology (GO) annotation and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis [29]. The GO categories encompassed three sections: biological processes (BP), molecular functions (MF), and cellular components (CC), with a significance threshold set at an adjusted P-value < 0.05 for the enrichment pathways. GSCALite database (http://bioinfo.life.hust.edu.cn/web/GSCALite/) was employed to investigate the cancer pathway activity associated with candidate BM genes. Additionally, the EMTome database (http://www.emtome.org/) was used to analyze the correlation between gene expression and EMT signature in pan-cancer. The GeneMANIA (http://www.genemania.org) was employed as a flexible online resource to visualize the interaction network of key BM genes and predict gene function and interactions [30].

Drug sensitivity analysis
The "calcPhenotype" function of the "oncoPredict" R package was selected to calculate predicted sensitivity values for each tumor sample, leveraging the training drug sensitivity data sourced from the Genomics of Drug Sensitivity in Cancer database and training expression data from the Cancer Therapeutics Response Portal [31]. Tumor samples were stratified into high and low expression groups based on the median expression level of each gene. A sensitivity score was used to assess drug responsiveness between these groups, where a lower score indicated a higher sensitivity to potential drugs.

Immune cell infiltration analysis
To evaluate the correlation between BM genes and the tumor immune microenvironment, we employed the R package "estimate" to calculate the stromal score, immune score, and estimate score for each ESCC patient [32]. The CIBERSORT algorithm was used to estimate the abundance of 22 distinct immune cells based on gene expression profiles [33]. The MCPcounter algorithm was applied to quantify the relative abundance of eight immune cells and two stromal cells within a tissue sample derived from transcriptomic data [34]. In addition, we also examined the association between BM genes and 40 immune checkpoint genes. In this study, the expression profile of candidate genes in patients was acquired from GSE53625.

Consensus clustering analysis
The unsupervised consensus clustering approach was applied to divide ESCC patients into the optimal number of clusters based on BM gene expression, utilizing the R package "ConsensusClusterPlus" [35]. Subsequently,

Zhang *et al. Biology Direct*   (2024) 19:79

Page 4 of 20

Gene Set Variation Analysis (GSVA) analysis was conducted with the gene set "c5.all.v2023.2.Hs.symbols" as input. GSVA scores were calculated using the R package limma to identify differentially enriched GO terms between the two subtypes, with $|t \text{ values}| > 2$ serving as the threshold for significance.

Cell culture
The human normal esophageal epithelial cell line HEEC and ESCC cell line TE-1 used in this study were obtained from the Cell Bank of the Chinese Academy of Sciences (Shanghai, China). The cells were cultured in Dulbecco's modified eagle medium (DMEM, ThermoFisher Scientific, USA) and RPMI-1640 medium (Gibco, USA), respectively. Both mediums were supplemented with 10%

fetal bovine serum (FBS, Gibco, USA) and maintained at a temperature of 37°C in a humidified environment with 5% CO_2. Additionally, the cells were regularly checked for mycoplasma contamination as needed.

Cell transfection and drug treatment
Two small interfering RNA (siRNA) sequences targeting human BGN were synthesized by Hanbio Biotechnology (Shanghai, China). A scrambled siRNA was also synthesized as a control. The plasmids were then transfected into the TE-1 cell line using Lipofectamine®3000 (ThermoFisher Scientific, USA) according to the manufacturer's instructions and incubated at 37 °C for 24 h. Docetaxel, paclitaxel, oxaliplatin and OSI-027 were purchased from MCE, USA.

Fig. 1 Single-cell analysis revealed abnormally elevated BM-related module scores in ESCC cells. (**A**) The scRNA-seq data was analyzed using the UMAP algorithm, resulting in eight cell types. (**B**) The identification and classification of cell types were performed utilizing a set of specific cell markers that enabled accurate annotation of each cell type. (**C**) The BM-related module score was calculated using the AUCell algorithm. (**D**) Computation of the BM-related module score utilizing the ssGSEA algorithm. ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$, ns, no significance

Zhang *et al. Biology Direct*   (2024) 19:79

Page 5 of 20

Western blot analysis

The cells were treated with ice-cold RIPA lysis buffer (Beyotime, China) to extract total protein. The protein concentration was measured using the BCA Protein Assay Kit (Beyotime Biotechnology, Shanghai, China). Subsequently, the proteins were separated by 10% sodium dodecyl sulphate polyacrylamide gel electrophoresis (PAGE) and transferred onto a polyvinylidene difluoride (PVDF) membrane (Millipore, USA). After blocking with 5% (w/v) non-fat milk for one hour at 25 °C, the PVDF membranes were incubated overnight at 4 °C with appropriately diluted primary antibodies against BGN (ab109369, 1: 1000 dilution, Abcam, USA), GAPDH (60004–1, 1:3000 dilution, Proteintech, China), N-cad-herin (4061S, 1: 1000 dilution, CST, USA) and E-cadherin (ab314063, 1: 1000 dilution, Abcam, USA), respectively. The PVDF membranes were then washed several times with TBST buffer (10 mM Tris–HCl pH 8.0, 150 mM NaCl, 0.05% Tween 20) and incubated with diluted HRP-conjugated IgG secondary antibodies (ab205719, 1: 10,000 dilution, Abcam, USA) for two hours at 25 °C. Finally, Protein bands were detected using enhanced chemiluminescence detection reagents (Thermo Pierce, Cramlington, UK) and quantified utilizing the Bio-Rad Image Lab software.

Cell proliferation assay

Cell proliferation was assessed using the Cell Counting Kit-8 (CCK8) and plate colony formation assay. For the CCK8 assay, cells were seeded in 96-well plates at a volume of 100 μL per well. After incubation for one day, three days, five days and seven days, 10 μL of CCK8 (MCE, USA) was added to each well and incubated at 37 °C for two hours. The optical density (OD) values were then measured using a microplate reader (BioTek, USA) at 450 nm. For the colony formation assay, cells in the logarithmic growth phase were selected and dissociated with trypsin to obtain a single-cell suspension. These cells were then inoculated into a 12-well plate and cultured for 10 days, with regular monitoring of their growth. Then, all colonies were fixed using 4% paraformaldehyde, stained with crystal violet, and captured using a digital camera.

Wound-healing assay

The transfected TE-1 cells were seeded into six-well plates with a density of 1×10^6 cells per well and cultured for 24 h. Cells adhered and scratched with 200 μL pipette tips. Then, the cells were washed with PBS twice and cultured in a serum-free medium. Images were captured at 0 h and 24 h.

Transwell assay

The transfected TE-1 cells were seeded into the upper chamber of a 24-well Transwell plate (8 μm pore size, Coring, USA). The assays involved culturing the cells with 200 μL serum-free RPMI-1640 medium, while the lower chambers were supplemented with 600 μL of RPMI-1640 medium containing 20% FBS. After incubation, the migrated cells were fixed with 4% paraformaldehyde and stained with crystal violet. The cells above the membrane in the upper chamber were removed with cotton swabs. Finally, a quantitative analysis was conducted by counting the stained cells in three randomly selected fields.
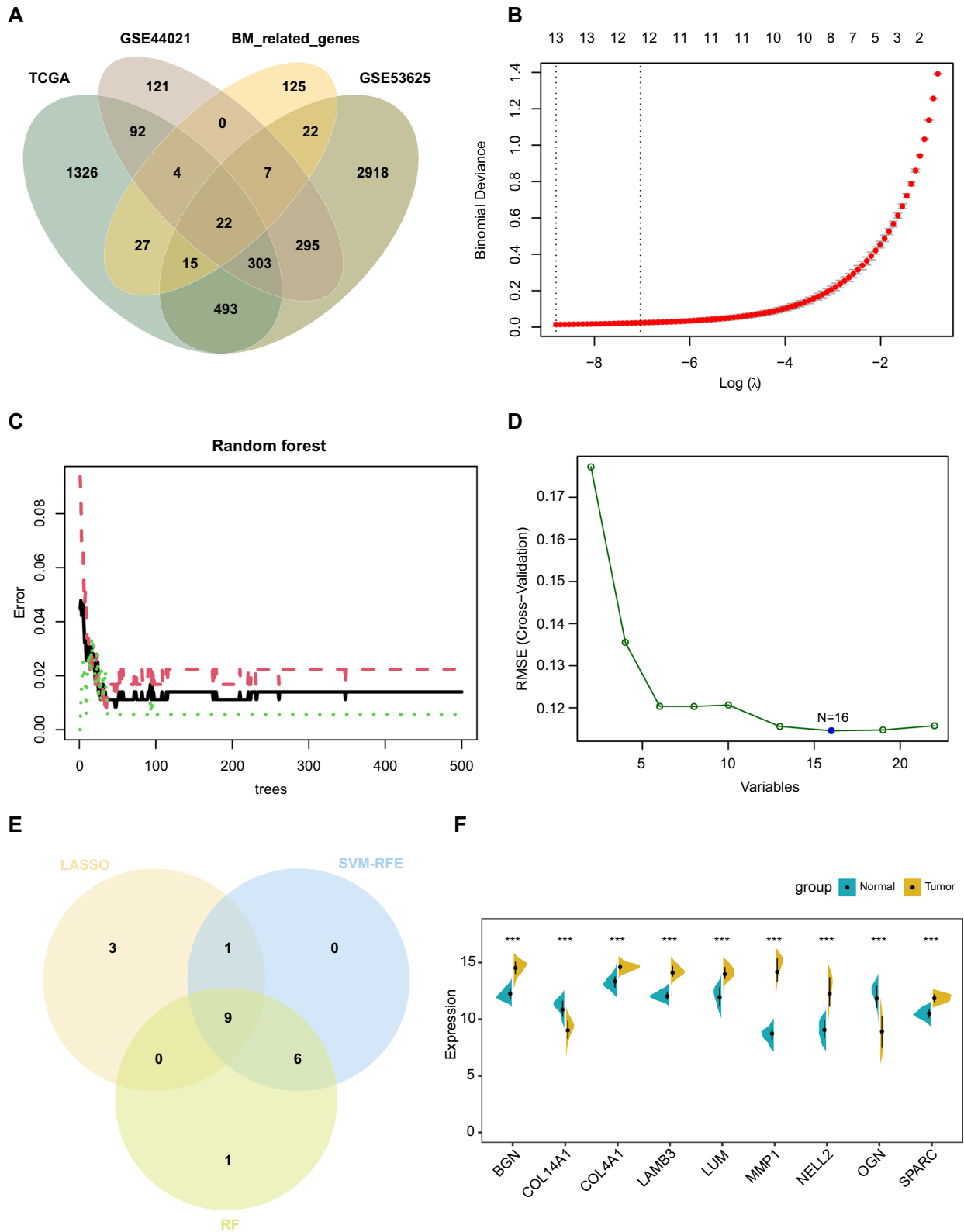
Flow cytometry

In brief, the cells were collected, washed, and re-suspended in a binding buffer to be monitored. Cells were processed with Annexin V-fluorescein isothiocyanate (FITC) and propidium iodide (PI) for apoptosis detection according to the protocol (Vazyme Biotech, China). Stained cells were detected and analyzed using a Canto II flow cytometer (BD Biosciences, Franklin Lakes, NJ, USA).

Statistical analysis

All experiments were conducted independently at least three times, yielding consistent outcomes, unless specified otherwise in the accompanying figure legends. All statistical analyses were performed using R software version 4.3.1. Comparisons between two different groups were carried out using the t-test. Correlation analysis was evaluated using Spearman's test. Statistical significance set at $P < 0.05$, ns > 0.05, and indicated by asterisks (* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, and **** $P < 0.0001$).

(See figure on next page.)
Fig. 2 Identification of BM-related diagnostic biomarkers for ESCC via machine learning methods. (**A**) Venn diagram showing the overlap between DEGs and 220 BMRGs. (**B**) Optimal λ values for hub genes identified by the LASSO logistic regression algorithm. (**C**) The correlation between the number of trees and the error rate in Random forests. (**D**) Gene selection using the SVM-RFE algorithm. (**E**) Venn diagram of diagnostic markers shared by LASSO, random forest, and SVM-RFE algorithms. (**F**) Split violin plot demonstrating the differential expression of nine hub genes between ESCC patients and controls in the GSE53625 dataset. *** $P < 0.001$

Zhang *et al. Biology Direct*   (2024) 19:79

Page 6 of 20

Fig. 2 (See legend on previous page.)

Zhang *et al. Biology Direct*   (2024) 19:79

Page 7 of 20

Results

Single-cell RNA-seq profiling demonstrates a significant correlation between ESCC occurrence and BMs

We analyzed scRNA-seq data using the UMAP method, revealing 18 distinct clusters. These clusters were meticulously classified into eight major cell types: T cells, B cells, myeloid cells, endothelial cells, epithelial cells, mast cells, fibroblasts, and smooth muscle cells (Fig. 1A). Expression profiles of marker genes for each cell type are depicted in the dotplot (Fig. 1B). Out of the 222 BM genes analyzed, 212 genes were identified in our single-cell investigation. Utilizing these genes, we calculated the BM-related module score. Analysis by AUCell (Fig. 1C) and ssGSEA (Fig. 1D) algorithms consistently indicated elevated BM scores in most ESCC cell types compared to normal cells, with the fibroblasts exhibiting the highest scores. This finding strongly suggests that a noteworthy association exists between the occurrence of ESCC and the emergence of BMs.

Machine learning-based screening of characteristic genes for clinical diagnosis of ESCC

In order to elucidate characteristic genes, we identified DEGs between the ESCC group and the control group from the TCGA, GSE53625, and GSE44021 datasets via the R package "limma". Venn diagrams revealed 22 genes shared by DEGs and 220 BMRGs (Fig. 2A). Subsequently, we conducted functional analysis on these 22 genes. GO analysis indicated their involvement in the extracellular matrix organization, collagen-containing extracellular matrix, basement membrane, and matrix structural constituent terms (Supplementary Fig. S1A-C). Additionally, KEGG analysis showed enrichment in ECM-receptor interaction, focal adhesion and PI3K-Akt signaling pathway (Supplementary Fig. S1D), with most genes enriched in basement membrane-related terms upregulated in ESCC group.

We further employed three machine learning algorithms to narrow down the diagnostic genes. The LASSO algorithm identified 13 characteristic genes according to the optimum λ value (Fig. 2B). Their coefficient profile is shown in Supplementary Fig. S1E. Using random forests, we optimized the model to achieve a minimum error rate (Fig. 2C) and selected genes with a Mean Decrease Gini (MDG) > 2 (Supplementary Fig. S1F), resulting in the

identification of 16 characteristic genes. Meanwhile, the SVM-RFE algorithm screened 16 genes with the minimum error (Fig. 2D). After intersecting the results from LASSO, RF and SVM-RFE, nine key diagnostic genes were identified (Fig. 2E). We examined the expression patterns of these nine genes between ESCC and control samples using the GSE53625 dataset (Fig. 2F). Finally, we selected seven upregulated hub genes in tumor tissues for further analysis, including *BGN, COL4A1, LAMB3, LUM, MMP1, NELL2,* and *SPARC.*
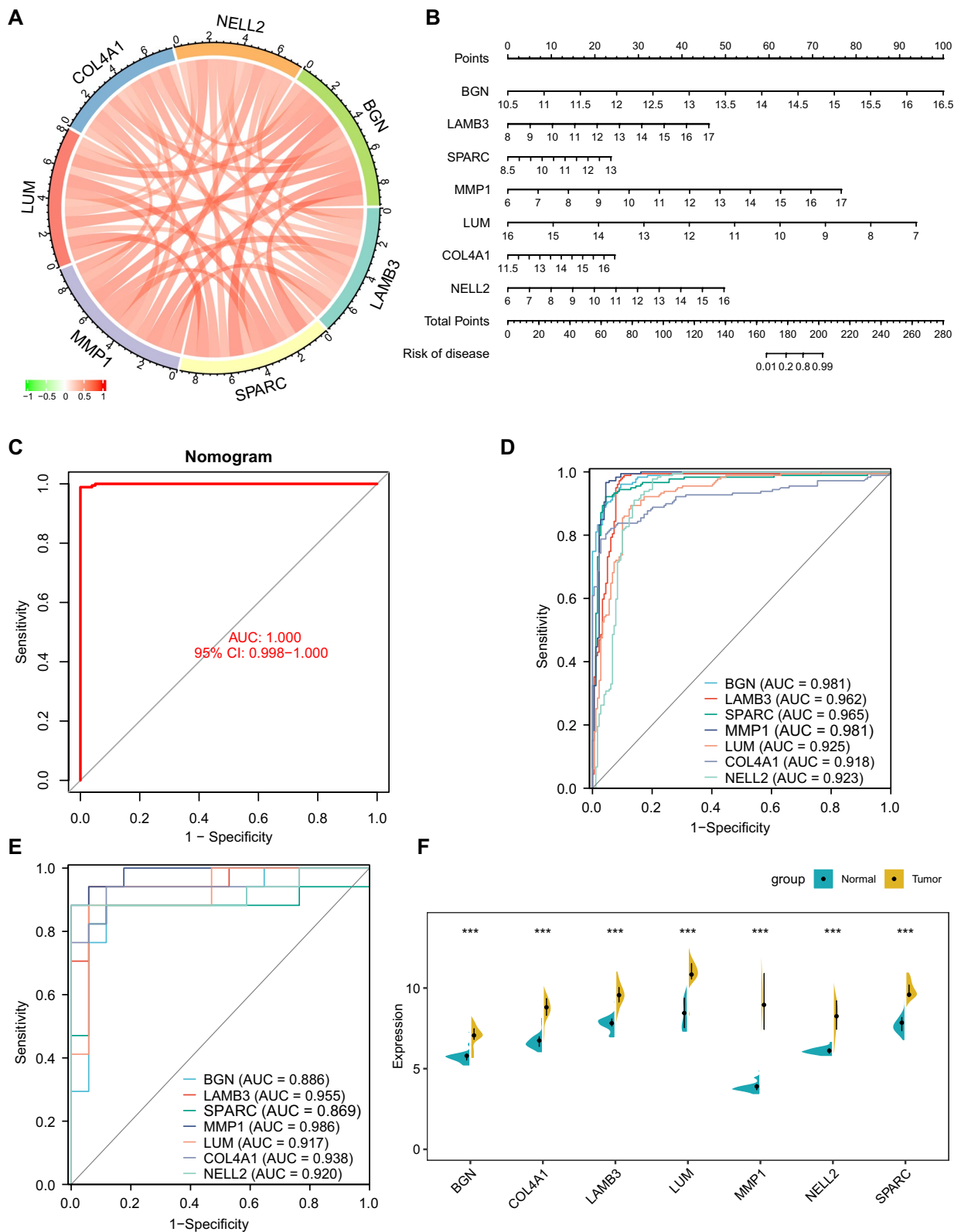
Validating potential biomarkers for ESCC diagnosis

In our study, it is evident that the seven selected genes demonstrate a distinct co-expression pattern in the GSE53625 dataset (Fig. 3A). Analysis using the Genemania network revealed strong interactions among these genes in various biological aspects (Supplementary Fig. S2A). Based on these seven genes, we constructed a nomogram exhibiting an excellent predictive performance with an AUC value of 1.0 (Fig. 3B, C). The calibration curves demonstrate that the diagnostic model of the constructed nomogram closely aligns with the ideal model, indicating a highly accurate predicted probability (Supplementary Fig. S2B). We further assessed the diagnostic efficacy of each gene in predicting the occurrence of ESCC. The results are summarized as follows: *BGN* (AUC 0.981, CI 0.967–0.992), *COL4A1* (AUC 0.918, CI 0.885–0.948), *LAMB3* (AUC 0.962, CI 0.939–0.981), *SPARC* (AUC 0.965, CI 0.943–0.984), *LUM* (AUC 0.925, CI 0.893–0.952), *MMP1* (AUC 0.981, CI 0.964–0.993), and *NELL2* (AUC 0.923, CI 0.891–0.953) (Fig. 3D).

To confirm the expression levels and diagnostic precision of these signature genes, we utilized two independent external datasets, GSE20347 and GSE23400. In the GSE20347 dataset, the AUC values for *BGN, LAMB3, SPARC, MMP1, LUM, COL4A1,* and *NELL2* were 0.886, 0.955, 0.869, 0.986, 0.917, 0.938, and 0.920, respectively (Fig. 3E). Moreover, the expression levels of these seven genes were significantly elevated in tumor samples compared to the control group (Fig. 3F). In the GSE23400 dataset, the AUC values for *BGN, LAMB3, SPARC, MMP1, LUM, COL4A1,* and *NELL2* were 0.921, 0.927, 0.859, 0.968, 0.827, 0.844, and 0.868, respectively (Supplementary Fig. S2C). Additionally, all seven genes

(See figure on next page.)

Fig. 3 Development and validation of a diagnostic nomogram model. (**A**) Correlation matrix of the seven gene expression profiles from the GSE53625 dataset, highlighting inter-gene relationships. (**B**) Construction of a predictive nomogram incorporating characteristic genes for ESCC. Each gene is assigned a score, with the total score representing the sum of individual gene scores. (**C**) ROC curve analysis of the nomogram model, demonstrating its diagnostic accuracy. (**D**) Comparative ROC curves for the individual predictive genes within the GSE53625 dataset. (**E**) ROC curves for the corresponding genes in the GSE20347 dataset illustrate their predictive performance. (**F**) Split violin plot contrasting gene expression levels between ESCC patients and controls within the GSE20347 dataset, underscoring the differential expression patterns. *** $P < 0.001$

Zhang *et al. Biology Direct*   (2024) 19:79

Page 8 of 20

Fig. 3 (See legend on previous page.)

Zhang *et al. Biology Direct*   (2024) 19:79

Page 9 of 20

Analysis of gene expression and tumor microenvironment in ESCC

Studies indicate that the tumor microenvironment modulates extracellular matrix components to foster tumorigenesis [36, 37]. We then conducted an in-depth analysis to elucidate the interplay between the tumor microenvironment and the seven BM gene expressions in ESCC patients. Utilizing a heatmap, we depicted the tumor microenvironment scores alongside the expression profiles of seven key BM genes, offering a visual synopsis of the molecular landscape in ESCC patients (Fig. 4A). Notably, the *BGN, MMP1, LUM* and *SPARC* genes exhibited positive associations with the immune score, while *COL4A1, LAMB3* and *NELL2* demonstrated negative correlations, suggesting a dichotomous role in immune modulation (Fig. 4B).

Further, we employed ssGSEA to scrutinize the variations in 13 immune-related functions between normal and tumor tissues. Our findings indicated that 11 of these immune function scores were significantly elevated in tumor tissues, underscoring the presence of an aberrant immune response in ESCC (Fig. 4C). To delve into the cellular composition of the immune response, we applied the CIBERSORT algorithm to quantify the abundance of 22 distinct immune cell populations. We discovered that M0 and M1 macrophages, CD4 memory-activated T cells, and activated dendritic cells were positively correlated with the expression of all seven BM genes. Conversely, Tregs, CD8 T cells, activated NK cells, and naive B cells exhibited negative correlations, highlighting a complex interplay between tumor-associated genes and immune cell dynamics (Fig. 4D).

Leveraging these insights, we performed an integrated analysis of the immunological profiles of candidate BMRGs. This analysis revealed significant positive correlations with a spectrum of immune checkpoint genes, particularly *CD276, CTLA4* and *TNFRSF* family members, indicating a potential immunological signature that may be harnessed for therapeutic development (Fig. 4E).

Molecular subtyping and immunophenotyping of ESCC reveal distinct tumor microenvironments

To delineate the molecular subtypes of ESCC and elucidate their associated molecular features, we further performed a cluster analysis using a consensus clustering approach based on the expression patterns of seven BMRGs derived from the GSE53625 dataset. The optimal number of clusters was determined to be two (k=2), which stratified 179 ESCC samples into two distinct subtypes: subtype A (n=54) and subtype B (n=125) (Fig. 5A). All the seven BMRGs elevated in subtype B compared to subtype A (Fig. 5B). To elucidate the biological differences between these subtypes, we performed GSVA analysis to assess GO enrichment. The results showed that subtype B exhibited upregulation of various pathways, including gap junction assembly, metalloendopeptidase activity, keratinocyte proliferation and migration, collagen metabolic process, collagen catabolic process, lamellipodium membrane and extracellular matrix disassembly (Fig. 5C, D). Considering the central role of basement membranes in the tumor microenvironment, we further evaluated the differences between the two subtypes. Subtype B was found to have elevated stromal score and immune scores compared to subtype A, suggesting a more complex and potentially aggressive tumor microenvironment (Fig. 5E). Additionally, we utilized the MCPcounter algorithm to quantify the relative proportions of different infiltrating immune and stromal cells. Notably, four infiltrating immune cell types — cytotoxic lymphocytes, NK cells, monocytic lineage cells and neutrophils — as well as two stromal cell types, endothelial cells, and fibroblasts, were found to be more abundant in subtype B. Among these, fibroblasts were identified as the predominant cell component, indicating their potential role in the tumorigenic process (Fig. 5F). This finding aligns with the scRNA-seq analysis, which reveals that fibroblasts possess the highest scores for BM-related gene modules in ESCC (Fig. 1C, D). These results suggest a significant association between fibroblasts and the pathophysiology of basement membrane lesions.

(See figure on next page.)

Fig. 4 Tumor microenvironment characterization and expression of seven BM genes in ESCC patients. (**A**) Heatmap representation displaying the tumor microenvironment score and the expression levels of the seven BM genes in ESCC patients. (**B**) Scatter plots illustrating the correlations between the expression of the seven BM genes and the immune score (left panel) and the estimate score (right panel), highlighting the association between gene expression and immune activity. (**C**) Bar graphs showing the differences in immune-related functional scores between normal and tumor groups, indicating the impact of the tumor microenvironment on immune response. (**D**) Correlation matrices displaying the relationships between the expression of the seven BM genes and 22 immune-related cell types, elucidating the interaction between these genes and the immune system. (**E**) Correlation matrices depicting the associations between the expression of the seven BM genes and immune checkpoint genes, revealing potential regulatory mechanisms in the tumor microenvironment. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$

Zhang *et al. Biology Direct*   (2024) 19:79

Page 10 of 20

Fig. 4 (See legend on previous page.)

Zhang *et al. Biology Direct*   (2024) 19:79

Page 11 of 20

EMT pathway activation and fibroblast heterogeneity in ESCC tumor progression

In our analysis of ESCC scRNA-seq data, we identified the expression of six BMRGs, with the LAMB3 gene notably absent from this dataset. Remarkably, the genes *BGN*, *LUM*, *SPARC*, and *COL4A1* exhibited high expression levels in fibroblastic and smooth muscle cells (Fig. 6A). This observation corroborated our previous cluster analysis and suggested a significant role for these BMRGs in the cancer-associated fibroblasts. To further elucidate the dynamics of fibroblast differentiation, we extracted fibroblasts to perform a pseudotime analysis using Monocle2, which delineated six distinct states of cell fate trajectories based on their gene expression patterns (Fig. 6B). The differential expression of *BGN*, *LUM*, *SPARC*, and *COL4A1* throughout fibroblast differentiation implies that these genes may exert distinct influences on tumor progression (Fig. 6C).

Expanding the scope of our investigation to a broader oncological context, we harnessed the GSCALite platform to explore the correlation between gene expression profiles and the activation of cancer-related signaling pathways in a comprehensive dataset encompassing 33 distinct malignancies from TCGA. Our analysis uncovered that *BGN* exhibited the most significant association with the activation of the EMT pathway among the seven evaluated genes, shedding light on the molecular mechanisms driving tumorigenesis (Fig. 6D). This finding was further substantiated by a focused analysis of the correlation between *BGN* gene expression and EMT in pan-cancer using the EMTome database, which aligned with the results from GSCALite (Fig. 6E). Recognizing the critical role of EMT in processes such as tumor initiation, progression, and metastasis, we resolved to investigate the specific function of BGN in ESCC.

BGN suppression attenuates the migratory capacity of ESCC cells and augments apoptotic activity
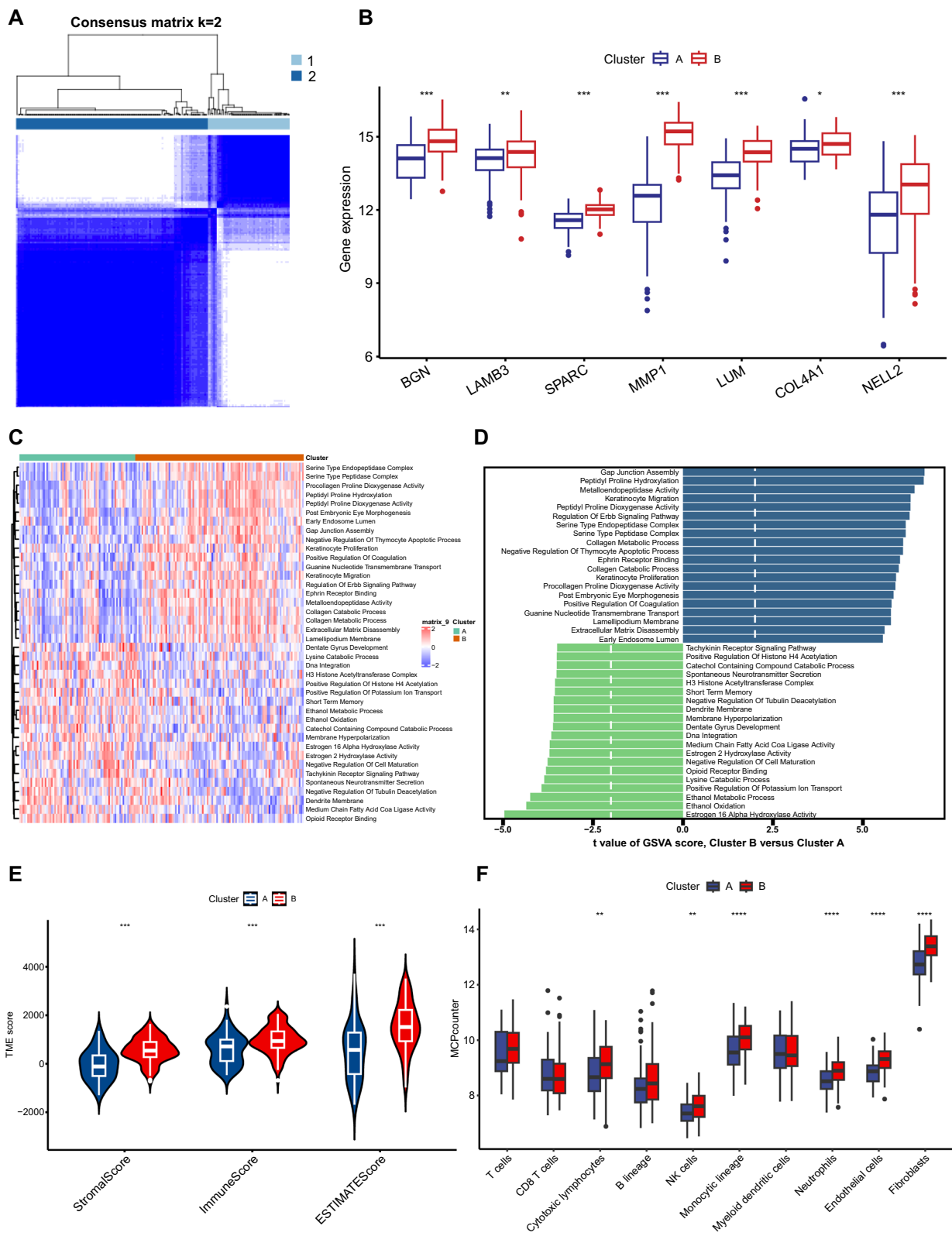
To investigate the role of *BGN* in ESCC pathophysiology, we utilized RNA interference to diminish endogenous BGN expression. Our preliminary assessment of BGN expression levels disclosed a markedly elevated expression in the TE-1 cell line when contrasted with that of HEEC cell line (Supplementary Fig. S3). Consequently, TE-1 was selected for subsequent investigation. Two distinct small interfering RNAs (siRNAs) and their corresponding scrambled controls were transfected into TE-1 cells, with knockdown efficacy confirmed by western blot analysis (Fig. 7A and Supplementary Fig. S4A). EMT, a process pivotally associated with cancer invasion and metastasis, involves transformating polarized epithelial cells into cells with mesenchymal phenotypes, characterized by the loss of cell adhesion and acquisition of migratory properties. Following BGN knockdown, there was a notable increase in the expression of the epithelial cell marker E-cadherin, coupled with a significant decrease in the mesenchymal cell marker N-cadherin (Fig. 7B and Supplementary Fig. S4B, C). Notably, BGN suppression led to a significant curtailment in cell proliferation and a reduction in clone formation in vitro (Fig. 7C, D). Flow cytometry analyses indicated that BGN knockdown enhanced apoptosis in TE-1 cells (Fig. 7E). Additionally, the migratory capabilities of TE-1 cells were substantially diminished upon BGN inhibition (Fig. 7F, G). Collectively, these findings underscore the tumor-suppressive effects of BGN inhibition in ESCC.

BGN, a member of the small leucine-rich proteoglycan (SLRP) family, class I subfamily, features 12 leucine-rich repeat (LRR) domains (Fig. 7H). AlphaFold's predicted structure of BGN (P21810) reveals an orderly arrangement of β-strands on the cell membrane, creating a 'palisade' configuration that likely mediates key protein interactions. Given the established role of glycosylation in protein function [38, 39], the O-linked glycosylation sites [40] at S42, S47, S180, S198, and N-linked sites at N270, N311 on BGN may critically influence its trafficking, localization, signaling, and immunological properties, providing a foundation for future research. Additionally, charge analysis identifies positively charged domains within the LRR ' palisade ' head (e.g., K139, K209, K238) and bottom regions (e.g., K82, K128, R266), which may serve as crucial sites for cofactor recruitment and protein interaction, suggesting potential targets for drug design (Fig. 7H).

(See figure on next page.)

Fig. 5 Characterization of molecular subtypes in ESCC and their associated molecular features. (**A**) Consensus clustering matrix demonstrating the optimal partitioning into two distinct subtypes when k = 2. (**B**) Box plots illustrating the mRNA expression levels of the seven BMRGs genes across the two identified subtypes. (**C**) Heatmap depicting the GSVA enrichment scores of GO pathways, highlighting the differences between the two subtypes. (**D**) Corresponding bar plot providing a detailed view of the GSVA enrichment differences in GO pathways between the subtypes. (**E**) Violin plot displaying the comparative stromal, immune, and estimate scores for the two subtypes, indicating variations in the tumor microenvironment. (**F**) Box plots showing the infiltration levels of eight immune cell types and two stromal cell types in the two subtypes, revealing the cellular composition of the tumor microenvironment * *P* < 0.05, ** *P* < 0.01, *** *P* < 0.001

Zhang *et al. Biology Direct*   (2024) 19:79

Page 12 of 20

Fig. 5 (See legend on previous page.)

Zhang *et al. Biology Direct*   (2024) 19:79

Page 13 of 20

BGN expression dictates sensitivity to chemotherapeutic agents in metastatic ESCC

In the therapeutic landscape of metastatic ESCC, chemotherapy regimens incorporating platinum and fluoropyrimidine/paclitaxel continue to set the standard of care [41]. Chemotherapy sensitivity refers to the extent of vulnerability of cancer cells to chemotherapeutic agents. Highly sensitive tumor cells are more likely to be eradicated during treatment, thereby enhancing the therapeutic efficacy of the drugs. Our approach commenced by stratifying ESCC patients into high and low-expression cohorts according to the median expression level of BGN. Subsequently, we assessed the differential drug sensitivity between these cohorts to identify candidate therapies for patients with elevated BGN expression. Notably, ESCC patients exhibiting high BGN expression demonstrated enhanced sensitivity to docetaxel, paclitaxel, and oxaliplatin, all established chemotherapeutic agents in ESCC management and mTOR inhibitor OSI-027 (Fig. 8A). To elucidate this, we conducted cell proliferation assays to assess the chemosensitivity of cells exhibiting varying levels of BGN expression to a panel of clinically relevant drugs. It is worth noting that the proliferation of TE-1 cells, which endogenously express high levels of BGN, was markedly suppressed upon treatment with docetaxel, paclitaxel, oxaliplatin and OSI-027, particularly when compared to cells with BGN knockdown (Fig. 8B–E). Altogether, our data suggest that BGN expression may serve as a predictive biomarker to guide clinical drug selection in ESCC treatment.

Discussion

ESCC is characterized by its aggressive phenotype, typically presenting at advanced stages with a pronounced tendency for metastasis and exhibiting resistance to chemotherapeutic agents. The BMs provide structural support and tissue compartmentalization, facilitating dynamic cellular interactions that regulate behaviors like proliferation and migration [13]. Additionally, BMs act as a physical barrier, typically impeding primary tumor dissemination during the in situ phase of cancer [42]. Meanwhile, the degradation of BMs is a critical initial step for ESCC cells to invade and metastasize. Therefore, investigating the early diagnosis and mechanisms of ESCC may enhance diagnostic precision and guide stratified therapeutics, addressing a critical research gap. With this in

mind, we aimed to investigate the clinical relevance of BM-associated genes in ESCC. Using single-cell data, we initially estimated the BM score in individual cells based on the expression profiles of related genes. Our integrative analysis identified a set of 22 differentially expressed genes enriched in BMs functions, which were further refined to seven key diagnostic genes (*BGN*, *COL4A1*, *LAMB3*, *LUM*, *MMP1*, *NELL2*, and *SPARC*) through machine learning algorithms. These genes showed distinct co-expression patterns and high diagnostic accuracy in ESCC, as validated across independent datasets, underscoring their potential as robust biomarkers for early detection and stratified treatment strategies. This approach holds the potential for diagnosing and treating a spectrum of solid tumors in future oncology applications.

Among the seven biomarkers, MMP-1, known for its BM-degrading capabilities, has been implicated as a potential diagnostic and prognostic biomarker in ESCC [43]. COL4A1, a key component of collagen IV, has been shown to promote tumorigenesis in various cancers and is consistently upregulated in ESCC [44]. Laminins, encoded by LAMB3, are glycoproteins integral to BM and have been linked to cancer cell proliferation and migration [45]. SPARC, a matricellular glycoprotein, is associated with aggressive metastatic tumors and is highly expressed in ESCC, where it may promote tumor progression through interactions with the tumor microenvironment [46]. LUM, a small leucine-rich proteoglycan, and NELL2, a secreted protein with potential roles in cancer development, were also identified as differentially expressed in ESCC. Biglycan (BGN) is a vital extracellular matrix constituent implicated in various malignancies [47–50]. It can be secreted by prostate cancer cells and regulate myeloid-derived suppressor cells (MDSCs) migration through Akt/mTOR and MNK/eIF4E pathways [51].

In the context of advanced cancers, tumor-associated macrophages (TAMs) are known to undermine immune surveillance by adopting immunosuppressive attributes that impede the antitumor efficacy of CD8$^+$ T cells. These TAMs exhibit an upregulation of PD-L1, which interfaces with PD-1 receptors on T cells, precipitating T cell exhaustion. Furthermore, they secrete IL-10, TGF-β, and arginase, as well as initiate collagen deposition and TME metabolic reprogramming, which collectively

(See figure on next page.)
Fig. 6 Molecular characterization and functional analysis of fibroblasts in ESCC. (**A**) Tissue-specific expression patterns of BMRGs including BGN, *MMP1, LUM, SPARC, COL4A1*, and *NELL2* across various cell types. (**B**) Pseudotime trajectory analysis of fibroblasts, illustrating differentiation stages (left) and progression along pseudotime (right). (**C**) Expression dynamics of BGN, COL4A1, LUM, and SPARC across different pseudotime points and differentiation stages. (**D**) The GSCALite platform was used to analyze the cancer-related pathway activity. (**E**) Examination of the correlation between *BGN* gene expression and EMT in pan-cancer using the EMTome database, revealing the widespread impact of BGN on cancer progression
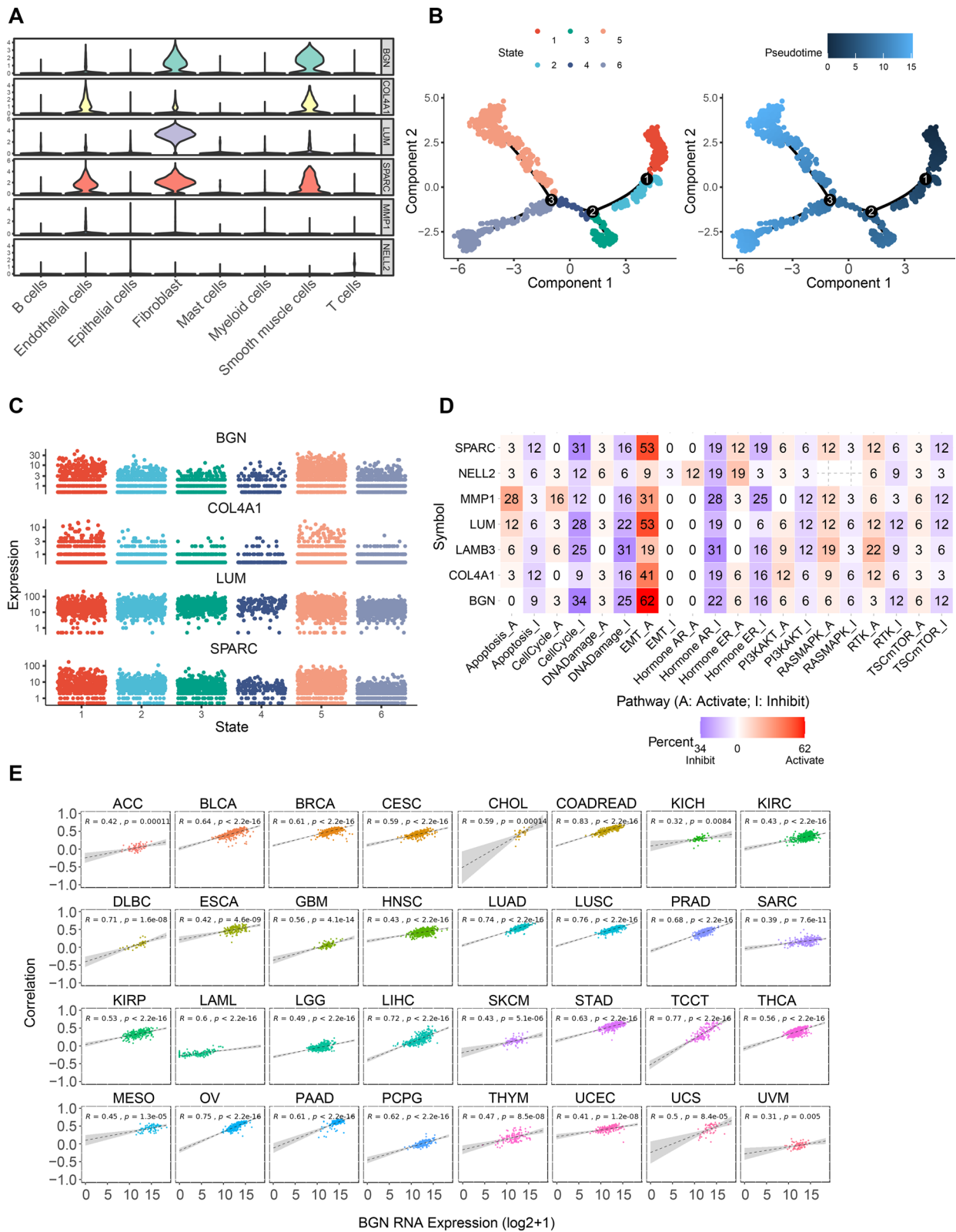
Fig. 6 (See legend on previous page.)

dampen the activity of CD8[+] T cells [52]. This immunosuppressive environment propels the differentiation of CD4[+] T cells into regulatory T cells (Tregs), thereby exacerbating the suppression of antitumor responses. Treg cells and TAMs are engaged in a reciprocal positive feedback mechanism that fosters an immunosuppressive phenotype [53]. This study revealed that the seven BM genes significantly correlate with M0 macrophages, M1 macrophages, and CD8 T cells, indicating their role in the tumor immune microenvironment (Fig. 4D). In addition, cancer-associated fibroblasts (CAFs) and TAMs interact through the SDF-1/CXCR4 signaling axis, leading to the activation of PI3K/Akt and NF-κB pathways, thereby reinforcing a sophisticated network of immune evasion [54, 55]. CAFs are critical in modulating the biomechanical properties of the ECM, contributing to its stiffness and degradation. These cells actively engage with a diverse array of cell types within the TME, thereby orchestrating ECM remodeling to facilitate the process of tumorigenesis [36]. In this work, our research delineates an ESCC tumor microenvironment dichotomy in BM gene expression, signifying immune response complexity and therapeutic targeting opportunities. ESCC molecular subtypes, distinguished by differential BMRG expression, reveal subtype B as more aggressive with enriched immune and stromal scores, highlighting the role of fibroblasts in tumorigenesis. BGN is hypothesized to enhance the migratory and invasive properties of malignant cancer cells [56, 57]. Its modulation by the TGF-β signaling pathway, a vital driver of EMT, highlights the importance of this interplay [58]. Furthermore, BGN has emerged as a biomarker that marks the progressive transition from normal colonic mucosa to adenoma and ultimately to adenocarcinoma, particularly in the context of colorectal cancer [59].

BGN's significant correlation with EMT pathway activation across cancers suggests its pivotal role in tumor progression, warranting further investigation into its specific function in ESCC (Fig. 6D, E). Although elevated BGN expression has been reported to correlate with poorer overall survival in ESCC, the underlying mechanisms remain to be elucidated [24]. In vitro experiments from our study corroborate these findings, showing that RNA interference-mediated BGN suppression in ESCC cell lines significantly impeded cell proliferation, induced apoptosis, and attenuated cell migration and invasion, revealing a tumor-suppressive role for BGN in ESCC pathophysiology.

Cancer represents a multifaceted systemic disorder characterized by the dynamic interplay between malignant cells, the ECM, and the diverse cellular constituents of the TME. The ECM's pivotal role in tumor progression marks it as a prime target for oncological therapies. TKIs have revolutionized NSCLC [60] and CML [61] treatment by targeting EGFR mutations, partially via discoidin domain receptors (DDRs) signal transduction inhibition. There is an urgent need to identify tumor-specific ECM targets that can suppress cell proliferation, migration, and angiogenesis to curb tumor progression without harming healthy tissues. Intratumoral ECM has been proposed as an antigen for tumor vaccines and CAR therapy, with studies demonstrating that fibronectin EDA domain-targeted vaccines improved macrophage infiltration, inhibited angiogenesis, and reduced metastasis in a breast cancer mouse model [62, 63]. Despite advances, challenges persist in ECM-targeted cancer therapies, notably the absence of systematic assessments of ECM variability across cancers, which is essential for tailored treatment strategies and sensitivity profiling. Furthermore, the scarcity of ECM-specific drug libraries may stem from inadequate in vitro and in vivo models [37]. Our stratification of ESCC patients by BGN expression levels identified a heightened sensitivity to standard chemotherapies in the high-expression cohort. TE-1 cells with intrinsic high BGN levels showed significantly reduced proliferation following treatment with docetaxel, paclitaxel, oxaliplatin and OSI-027, in accordance with the drug sensitivity analysis. These findings position BGN as a promising biomarker for personalized ESCC chemotherapy and highlight its potential role in guiding the selection of targeted therapeutics.

(See figure on next page.)

Fig. 7 Impact of BGN knockdown on ESCC cell phenotypes. (**A**) Western blot analysis confirming the significant suppression of BGN expression following siRNA transfection in ESCC cells. (**B**) Knockdown BGN led to a significant downregulation of the EMT marker N-cadherin, accompanied by an upregulation of E-cadherin. (**C**) CCK-8 assay depicting the attenuated proliferation of TE-1 cells upon BGN knockdown. (**D**) Plate cloning assay illustrating the reduced clone formation capacity of TE-1 cells with BGN depletion. (**E**) Flow cytometry analysis revealing the influence of BGN knockdown on apoptosis in TE-1 cells, focusing on the role of *BGN*. (**F-G**) Transwell assays (**F**) and wound-healing assays (**G**) quantifying the compromised migration capabilities of TE-1 cells post-BGN knockdown. Error bars, ± SD from at least three biological replications. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, ****$P < 0.0001$. (**H**) In the left panel, the AlphaFold-predicted structure of the BGN protein (P21810) is depicted in a cartoon representation, with O-linked glycosylation sites at S42, S47, S180 and S198, as well as N-linked glycosylation sites at N270 and N311, illustrated using a ball-and-stick model. The 12 leucine-rich repeat (LRR) domains are delineated within dashed squares. The right panel presents a rotated 180° view of the charge distribution, with elliptical dashed lines highlighting the positively charged regions at the head and tail of the 'palisade' structure

Zhang *et al. Biology Direct*   (2024) 19:79

Page 16 of 20
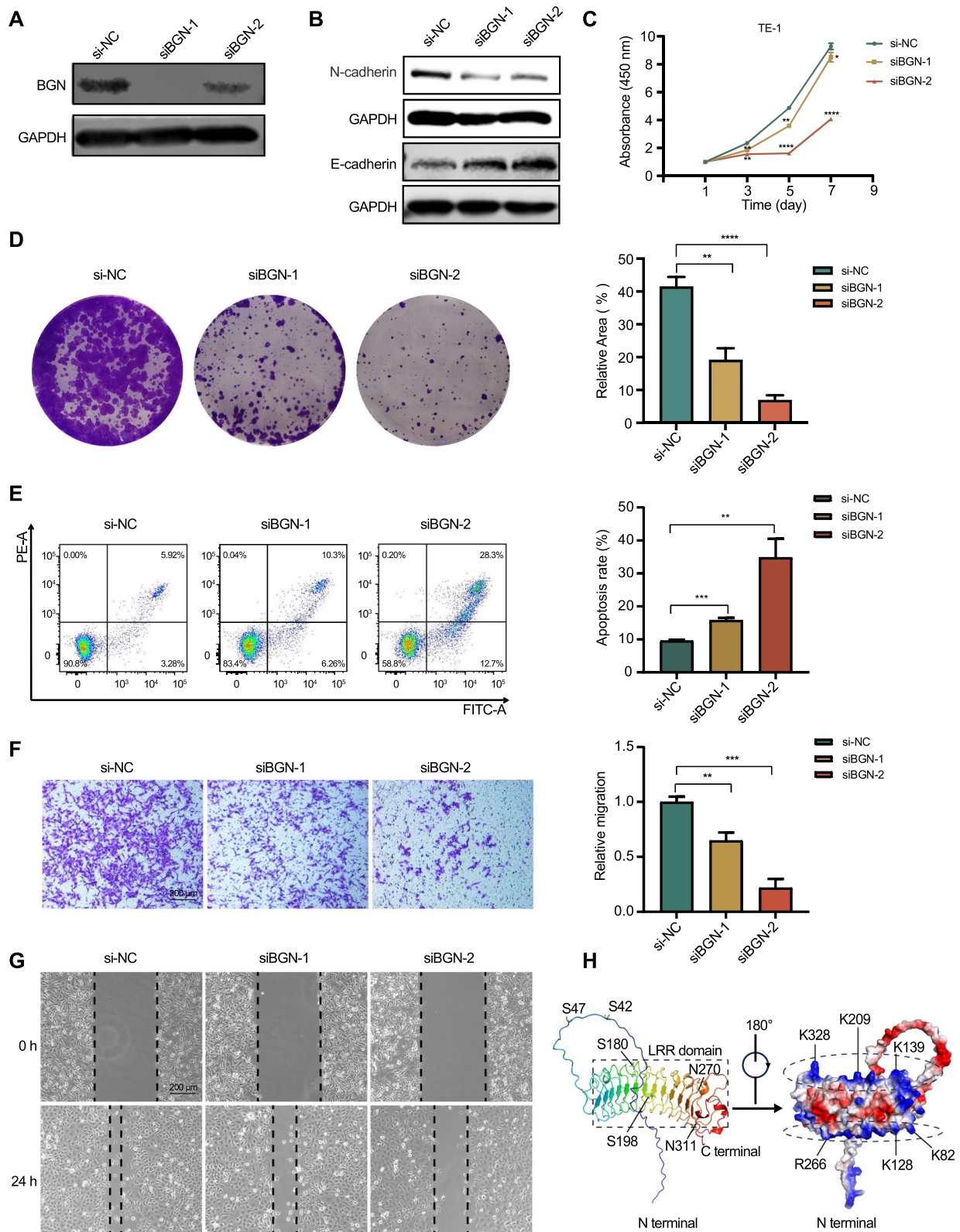
Fig. 7 (See legend on previous page.)
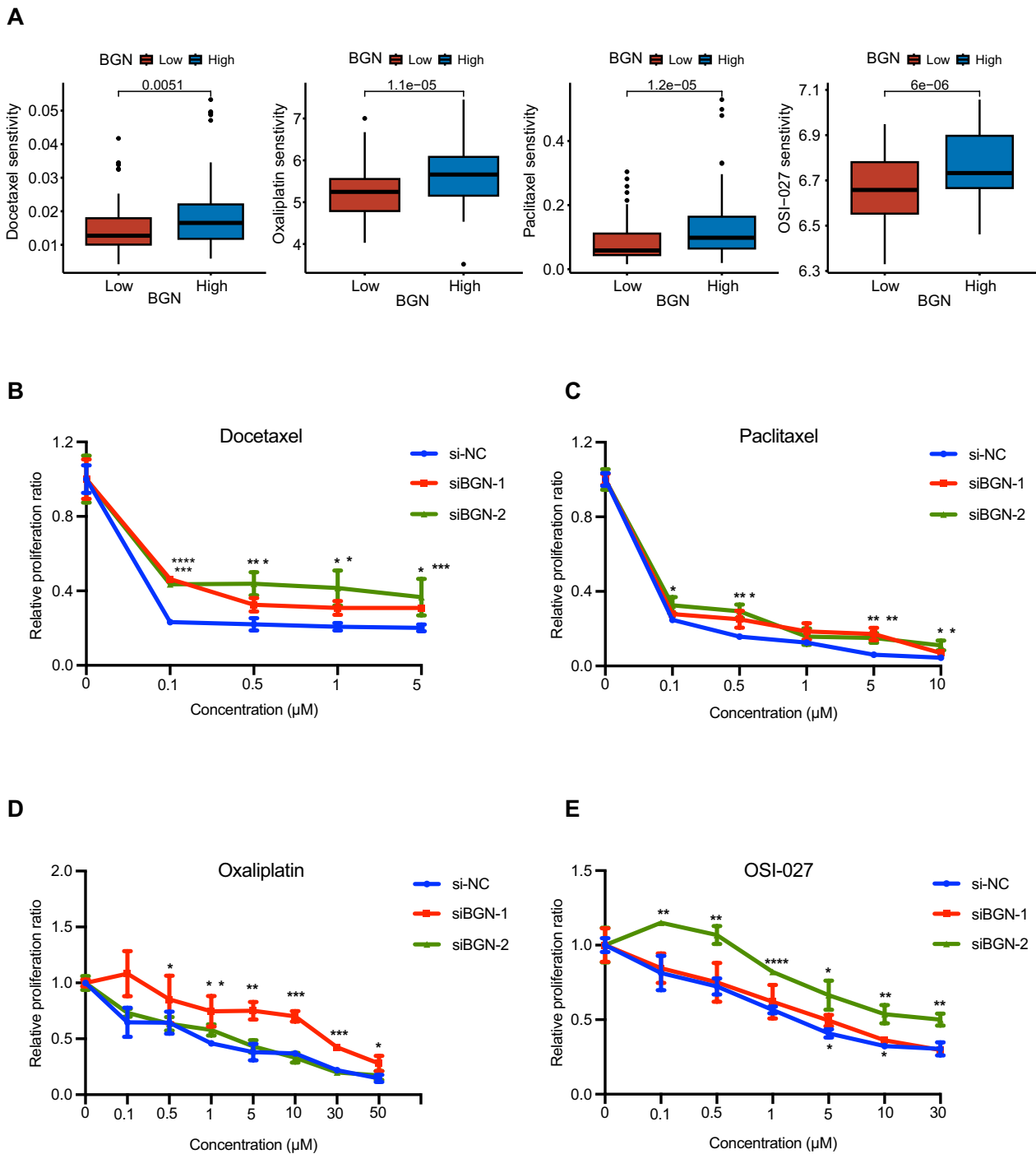
Fig. 8 BGN expression and its correlation with drug sensitivity in ESCC. (**A**) Predicted sensitivity profiles for the chemotherapeutic agents docetaxel, paclitaxel, oxaliplatin and OSI-027, stratified by BGN expression levels. (**B-E**) Proliferation assays showing the effects of docetaxel (**B**), paclitaxel (**C**), oxaliplatin (**D**) and OSI-027 (**E**) drugs on BGN knockdown TE-1 cells, highlighting the influence of BGN expression on treatment response, suggesting a potential role of BGN in dictating therapeutic susceptibility. Error bars, ± SD from three biological replications. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$

Zhang *et al. Biology Direct*   (2024) 19:79

Page 18 of 20

Conclusion

In conclusion, our research has pinpointed seven pivotal characteristic genes and crafted a diagnostic nomogram integrating these genetic markers with clinical metrics, offering a robust framework for the early detection of ESCC. These genetic elements may also serve as prognostic indicators and therapeutic targets, facilitating personalized clinical strategies and medical decision-making for ESCC management. Given BGN's crucial involvement in the etiology and progression of ESCC, therapeutic targeting of this protein could confer a significant clinical benefit. Thus, the adjunctive use of chemotherapeutic agents like docetaxel, paclitaxel, and oxaliplatin, in tandem with current therapeutic strategies, may significantly bolster the clinical efficacy of ESCC treatments.

Abbreviations

AUC	Area under the curve
BM	Basement membrane
BMRGs	Basement membrane-related genes
DEGs	Differentially expressed genes
ECM	Extracellular matrix
EMT	Epithelial-mesenchymal transition
ESCC	Esophageal squamous cell carcinoma
GSVA	Gene Set Variation Analysis
KEGG	Kyoto Encyclopedia of Genes and Genomes
LASSO	The least absolute shrinkage and selection operator
RF	Random forest
ROC	Receiver operating characteristic curve
SVM-RFE	Support vector machine-recursive feature elimination
TME	Tumor microenvironment

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13062-024-00529-3.

Supplementary Material 1: Figure S1. Functional analysis reveals the involvement of BM-related genes in ESCC Tumors. (A-C) Bubble plots displaying main GO terms for BPs, CCs, and MFs associated with BM-related DEGs. (D) KEGG pathway enrichment analysis of the BM-related DEGs. (E) LASSO coefficients profile showing the hub genes. (F) The Random Forests method identifies characteristic genes with a Mean Decrease Gini value greater than two. GO, the Gene Ontology; BPs, Biological Processes, CCs, Cellular Components, MFs, Molecular Functions, KEGG, the Kyoto Encyclopedia of Genes and Genomes analyses.

Supplementary Material 2: Figure S2. Gene interaction networks and diagnostic assessment of ESCC. (A) Genemania-generated network diagram illustrating the interconnectivity among BGN, LAMB3, SPARC, MMP1, LUM, COL4A1, and NELL2, highlighting the complex relationships within the ESCC molecular landscape. (B) Calibration curve analysis, demonstrating the predictive accuracy of the nomogram across various probability thresholds. (C) ROC curves for the seven genes in the GSE23400 dataset, providing an assessment of their diagnostic efficacy. (D) Split violin plot delineating the gene expression disparities between ESCC patients and controls in the GSE23400 dataset, emphasizing the distinct transcriptional profiles. *** $P < 0.001$.

Supplementary Material 3: Figure S3. Western blot analysis demonstrated a substantial upregulation of BGN expression in TE-1 cells relative to HEEC cells. (A-B) Quantitative western blot analysis, performed with replicates exceeding the triplicate standard, indicates a statistically significant increase in BGN expression levels within TE-1 cells in contrast to those in HEEC cells. ** $P < 0.01$.

Supplementary Material 4: Figure S4. Quantitative analysis of protein expression by western blot. (A) Quantitative assessment of BGN expression was performed with a minimum of three independent biological replicates. (B-C) Statistical quantification of N-cadherinand E-cadherinexpression was conducted with a minimum of three biological replicates. * $P < 0.05$, ** $P < 0.01$.

Supplementary Material 5: Table S1. BM-related genes.

Author contributions
Conceived and designed the study: HZ. Performed the experiments: XZ, ZYH. Analyzed the data: HZ and XZ. Wrote the original manuscript: HZ. Revised the manuscript: HZ. Project supervision: HZ.All authors reviewed the manuscript.

Availability of data and materials
The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate
This study was approved by Ethics Committee of Shanghai Chest Hospital, Shanghai Jiao Tong University School of Medicine.

Consent for publication
This study has not been published before, and all authors approved this publication.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Department of Central Laboratory, Shanghai Chest Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200030, People's Republic of China. [2]Department of Geriatrics, Medical Center On Aging of Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200025, People's Republic of China.

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.
2. Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, et al. Cancer statistics in China, 2015. CA Cancer J Clin. 2016;66:115–32.
3. Miller KD, Nogueira L, Mariotto AB, Rowland JH, Yabroff KR, Alfano CM, et al. Cancer treatment and survivorship statistics, 2019. CA Cancer J Clin. 2019;69:363–85.
4. Huang J, Koulaouzidis A, Marlicz W, Lok V, Chu C, Ngai CH, et al. Global burden, risk factors, and trends of esophageal cancer: an analysis of cancer registries from 48 Countries. Cancers (Basel). 2021;13:141.
5. Welch DR, Hurst DR. Defining the hallmarks of metastasis. Cancer Res. 2019;79:3011–27.
6. Hall A. The cytoskeleton and cancer. Cancer Metastasis Rev. 2009;28:5–14.
7. Hanahan D, Weinberg RA. The hallmarks of cancer. Cell. 2000;100:57–70.
8. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011;144:646–74.

Zhang *et al. Biology Direct*   (2024) 19:79

Page 19 of 20

9.   Reuten R, Zendehroud S, Nicolau M, Fleischhauer L, Laitala A, Kiderlen S, et al. Basement membrane stiffness determines metastases formation. Nat Mater. 2021;20:892–903.

10. Wolf K, Wu YI, Liu Y, Geiger J, Tam E, Overall C, et al. Multi-step pericellular proteolysis controls the transition from individual to collective cancer cell invasion. Nat Cell Biol. 2007;9:893–904.

11. Reuter JA, Ortiz-Urda S, Kretz M, Garcia J, Scholl FA, Pasmooij AM, et al. Modeling inducible human tissue neoplasia identifies an extracellular matrix interaction network involved in cancer progression. Cancer Cell. 2009;15:477–88.

12. LeBleu VS, Macdonald B, Kalluri R. Structure and function of basement membranes. Exp Biol Med (Maywood). 2007;232:1121–9.

13. Yurchenco PD. Basement membranes: cell scaffoldings and signaling platforms. Cold Spring Harb Perspect Biol. 2011;3:a004911.

14. Pozzi A, Yurchenco PD, Iozzo RV. The nature and biology of basement membranes. Matrix Biol. 2017;57–58:1–11.

15. Strongin AY. Mislocalization and unconventional functions of cellular MMPs in cancer. Cancer Metastasis Rev. 2006;25:87–98.

16. Tian W, Luo Y, Tang Y, Kong Y, Wu L, Zheng S, et al. Novel implication of the basement membrane for breast cancer outcome and immune infiltration. Int J Biol Sci. 2023;19:1645–63.

17. Zhang Z, Zhu H, Wang X, Lin S, Ruan C, Wang Q. A novel basement membrane-related gene signature for prognosis of lung adenocarcinomas. Comput Biol Med. 2023;154: 106597.

18. Feng L, Yang J, Zhang W, Wang X, Li L, Peng M, et al. Prognostic significance and identification of basement membrane-associated lncRNA in bladder cancer. Front Oncol. 2022;12: 994703.

19. Tao J, Li X, Liang C, Liu Y, Zhou J. Expression of basement membrane genes and their prognostic significance in clear cell renal cell carcinoma patients. Front Oncol. 2022;12:1026331.

20. Wang GQ, Abnet CC, Shen Q, Lewin KJ, Sun XD, Roth MJ, et al. Histological precursors of oesophageal squamous cell carcinoma: results from a 13 year prospective follow up study in a high risk population. Gut. 2005;54:187–92.

21. Thiery JP, Acloque H, Huang RY, Nieto MA. Epithelial-mesenchymal transitions in development and disease. Cell. 2009;139:871–90.

22. Chen L, Zhu S, Liu T, Zhao X, Xiang T, Hu X, et al. Aberrant epithelial cell interaction promotes esophageal squamous-cell carcinoma development and progression. Signal Transduct Target Ther. 2023;8:453.

23. Zou B, Guo D, Kong P, Wang Y, Cheng X, Cui Y. Integrative genomic analyses of 1145 patient samples reveal new biomarkers in esophageal squamous cell carcinoma. Front Mol Biosci. 2021;8: 792779.

24. Tang L, Chen Y, Peng X, Zhou Y, Jiang H, Wang G, et al. Identification and validation of potential pathogenic genes and prognostic markers in ESCC by integrated bioinformatics analysis. Front Genet. 2020;11: 521004.

25. Jayadev R, Morais M, Ellingford JM, Srinivasan S, Naylor RW, Lawless C, et al. A basement membrane discovery pipeline uncovers network complexity, regulators, and human disease associations. Sci Adv. 2022;8:eabn2265.

26. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. J Stat Softw. 2010;33:1–22.

27. 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 Bioinf. 2018;19:432.

28. Ishwaran H, Kogalur UB. Consistency of Random Survival Forests. Stat Probab Lett. 2010;80:1056–64.

29. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS. 2012;16:284–7.

30. Warde-Farley D, Donaldson SL, Comes O, Zuberi K, Badrawi R, Chao P, et al. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. Nucleic Acids Res. 2010;38:W214–20.

31. Maeser D, Gruener RF, Huang RS. oncoPredict: an R package for predicting in vivo or cancer patient drug response and biomarkers from cell line screening data. Brief Bioinform. 2021. https://doi.org/10.1093/bib/bbab2 60.

32. Yoshihara K, Shahmoradgoli M, Martinez E, Vegesna R, Kim H, Torres-Garcia W, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. Nat Commun. 2013;4:2612.

33. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods. 2015;12:453–7.

34. Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. Genome Biol. 2016;17:218.

35. Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. Bioinformatics. 2010;26:1572–3.

36. Najafi M, Farhood B, Mortezaee K. Extracellular matrix (ECM) stiffness and degradation as cancer drivers. J Cell Biochem. 2019;120:2782–90.

37. Huang J, Zhang L, Wan D, Zhou L, Zheng S, Lin S, et al. Extracellular matrix and its therapeutic potential for cancer treatment. Signal Transduct Target Ther. 2021;6:153.

38. Hewa Nadugala B, Hantink R, Nebl T, White J, Pagel CN, Ranadheera CS, et al. The role of glycosylation in amyloid fibril formation of bovine kappa-casein. Curr Res Food Sci. 2023;6: 100433.

39. Jiang N, Chen W, Jothikumar P, Patel JM, Shashidharamurthy R, Selvaraj P, et al. Effects of anchor structure and glycosylation of Fcgamma receptor III on ligand binding affinity. Mol Biol Cell. 2016;27:3449–58.

40. Toledo AG, Pihl J, Spliid CB, Persson A, Nilsson J, Pereira MA, et al. An affinity chromatography and glycoproteomics workflow to profile the chondroitin sulfate proteoglycans that interact with malarial VAR2CSA in the placenta and in cancer. Glycobiology. 2020;30:989–1002.

41. Lordick F, Mariette C, Haustermans K, Obermannova R, Arnold D, Committee EG. Oesophageal cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2016;27:v50–7.

42. Kelley LC, Lohmer LL, Hagedorn EJ, Sherwood DR. Traversing the basement membrane in vivo: a diversity of strategies. J Cell Biol. 2014;204:291–302.

43. Zhang J, Zhang N, Yang X, Xin X, Jia CH, Li S, et al. Machine learning and novel biomarkers associated with immune infiltration for the diagnosis of esophageal squamous cell carcinoma. J Oncol. 2022;2022:6732780.

44. Wang T, Jin H, Hu J, Li X, Ruan H, Xu H, et al. COL4A1 promotes the growth and metastasis of hepatocellular carcinoma cells by activating FAK-Src signaling. J Exp Clin Cancer Res. 2020;39:148.

45. Zhang H, Pan YZ, Cheung M, Cao M, Yu C, Chen L, et al. LAMB3 mediates apoptotic, proliferative, invasive, and metastatic behaviors in pancreatic cancer by regulating the PI3K/Akt signaling pathway. Cell Death Dis. 2019;10:230.

46. Takayama H, Kobayashi S, Gotoh K, Sasaki K, Iwagami Y, Yamada D, et al. SPARC accelerates biliary tract cancer progression through CTGF-mediated tumor-stroma interactions: SPARC as a prognostic marker of survival after neoadjuvant therapy. J Cancer Res Clin Oncol. 2023;149:10935–50.

47. Wu H, Xiang Z, Huang G, He Q, Song J, Dou R, 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.

48. Zheng S, Zou Y, Tang Y, Yang A, Liang JY, Wu L, et al. Landscape of cancer-associated fibroblasts identifies the secreted biglycan as a protumor and immunosuppressive factor in triple-negative breast cancer. Oncoimmunology. 2022;11:2020984.

49. Liu Y, Chen S, Wang Y, Zhang Z, Wang Z, Tao Z, et al. The diagnostic and prognostic value of the immune checkpoint BGN in thymoma. Biochem Genet. 2023;62:1872.

50. Hu S, Xiao Q, Gao R, Qin J, Nie J, Chen Y, et al. Identification of BGN positive fibroblasts as a driving factor for colorectal cancer and development of its related prognostic model combined with machine learning. BMC Cancer. 2024;24:516.

51. Brina D, Ponzoni A, Troiani M, Cali B, Pasquini E, Attanasio G, et al. The Akt/mTOR and MNK/eIF4E pathways rewire the prostate cancer translatome to secrete HGF, SPP1 and BGN and recruit suppressive myeloid cells. Nat Cancer. 2023;4:1102–21.

52. Tharp KM, Kersten K, Maller O, Timblin GA, Stashko C, Canale FP, et al. Tumor-associated macrophages restrict CD8(+) T cell function through collagen deposition and metabolic reprogramming of the breast cancer microenvironment. Nat Cancer. 2024;5:1045.

53. Sun W, Wei FQ, Li WJ, Wei JW, Zhong H, Wen YH, et al. A positive-feedback loop between tumour infiltrating activated Treg cells and type 2-skewed

Zhang *et al. Biology Direct*   (2024) 19:79

Page 20 of 20

macrophages is essential for progression of laryngeal squamous cell carcinoma. Br J Cancer. 2017;117:1631–43.

54. Song T, Dou C, Jia Y, Tu K, Zheng X. TIMP-1 activated carcinoma-associated fibroblasts inhibit tumor apoptosis by activating SDF1/CXCR4 signaling in hepatocellular carcinoma. Oncotarget. 2015;6:12061–79.
55. Kloosterman DJ, Akkari L. Macrophages at the interface of the co-evolving cancer ecosystem. Cell. 2023;186:1627–51.
56. Sun H, Wang X, Zhang Y, Che X, Liu Z, Zhang L, et al. Biglycan enhances the ability of migration and invasion in endometrial cancer. Arch Gynecol Obstet. 2016;293:429–38.
57. Gu X, Ma Y, Xiao J, Zheng H, Song C, Gong Y, et al. Up-regulated biglycan expression correlates with the malignancy in human colorectal cancers. Clin Exp Med. 2012;12:195–9.
58. Chen WB, Lenschow W, Tiede K, Fischer JW, Kalthoff H, Ungefroren H. Smad4/DPC4-dependent regulation of biglycan gene expression by transforming growth factor-beta in pancreatic tumor cells. J Biol Chem. 2002;277:36118–28.
59. Mikula M, Rubel T, Karczmarski J, Goryca K, Dadlez M, Ostrowski J. Integrating proteomic and transcriptomic high-throughput surveys for search of new biomarkers of colon tumors. Funct Integr Genomics. 2011;11:215–24.
60. Goldberg SB, Redman MW, Lilenbaum R, Politi K, Stinchcombe TE, Horn L, et al. Randomized trial of afatinib plus cetuximab versus afatinib alone for first-line treatment of EGFR-mutant non-small-cell lung cancer: final results from SWOG S1403. J Clin Oncol. 2020;38:4076–85.
61. Jabbour E, Kantarjian H. Chronic myeloid leukemia: 2012 update on diagnosis, monitoring, and management. Am J Hematol. 2012;87:1037–45.
62. Femel J, Huijbers EJ, Saupe F, Cedervall J, Zhang L, Roswall P, et al. Therapeutic vaccination against fibronectin ED-A attenuates progression of metastatic breast cancer. Oncotarget. 2014;5:12418–27.
63. Lin EY, Jones JG, Li P, Zhu L, Whitney KD, Muller WJ, et al. Progression to malignancy in the polyoma middle T oncoprotein mouse breast cancer model provides a reliable model for human diseases. Am J Pathol. 2003;163:2113–26.

Publisher's Note