Submitted: 05 June, 2023 Accepted: 12 July, 2023 Published: 30 March, 2024 DOI:10.22514/jomh.2024.036

ORIGINAL RESEARCH

JOMH Journal of Men's Health

Exploring the significance of glycosylation in prostate cancer subtyping through single-cell analysis

Shijun Tong[1], Wenhui Zhu[1], Jing Zhai[1], Guanxiong Ding[1,*]

[1]Department of Urology, Huashan Hospital, Fudan University, 200040 Shanghai, China

***Correspondence**
10111220030@fudan.edu.cn
(Guanxiong Ding)

Abstract

Prostate cancer impacts millions of men worldwide and causes significant disease burden. Glycosylation is the post-translational modification offering novel therapeutics for prostate cancer. The scRNA-seq data is combined with bulk RNA-seq data of prostate cancer to understand the glycosylation role and identify the therapeutic targets. This study aims to investigate the differences within tumor and the role of glycosylation. The findings confirm that glycosylation can establish multiple cell biomarkers and divide the cell subtypes of prostate cancer. The specific cell subtypes have diverse functions in cell interactions, transcript activity, prognosis and immunotherapy response, such as UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 7+ (*GALNT7*+) epithelial cells and UDP Glucose Ceramide Glucosyltransferase+ (*UGCG*+) cancer associated macrophages. These outcomes assist in the better understanding of prostate cancer and provide new approach of the targeted therapy.

Keywords

Prostate cancer; Glycosylation; Single-cell analysis; Subtyping

1. Introduction

Prostate cancer (PCa) is the globally prevalent disease affecting millions of men. The incidence and mortality rates are high despite the progress in its diagnosis and treatment. Around one quarter of new male cancers diagnosed in United States are attributed to PCa [1]. Various treatments have been identified in recent years, including androgen receptor (AR) targeted therapies, chemotherapy, bone-targeting treatments, poly (adenosine diphosphate (ADP)-ribose) polymerase inhibitors and immunotherapy [2]. Immunotherapy has been particularly focused in PCa treatment. Other strategies include the targeting of cytotoxic T-lymphocyte associated antigen 4 (*CTLA4*), programmed cell death 1 (*PD1*) and programmed cell death ligand 1 (*PD-L1*), however their benefits for PCa patients are unclear. Studies have shown the potential of immunotherapy in treating PCa but with limited efficacy [3, 4]. There is a need to dissect the tumor microenvironment (TME) and discover potential biomarkers for improving the immunotherapeutic response in PCa patients.

Glycosylation involves the addition of sugar molecules to proteins, lipids or other biomolecules [5]. Glycosylation is the post-translational modification (PTM) occurring on range of cellular proteins [6]. The glycome includes diverse and abundant glycans on cells surface [7, 8]. Ten types of monosaccharides construct glycans in human which include glucose (Glc), galactose (Gal), N-acetylglucosamine (GlcNAc), N-acetylgalactosamine (GalNAc), fucose (Fuc), xylose (Xyl), sialic acid (Neu5Ac), glucuronic acid (GlcA), mannose (Man) and iduronic acid (IdoA) [5]. Glycosylation as an important regulatory mechanism controls multiple biological processes such as regulating the localization, function and activity of proteins in cells and tissues, and affects the biological activities of cell recognition, differentiation, signal transduction and immune response [5, 9, 10].

Abnormal glycosylation may cause diseases like cancer, diabetes, cardiovascular and neurodegenerative [11]. Glycosylation modification has role in cancer biology including the PCa [12]. N-Glycosylation is highly related to the clinical pathological changes of PCa [13, 14]. Recent research has identified specific changes in the glycosylation patterns of proteins in PCa that may serve as potential disease biomarkers. One type of glycosylation implicated in PCa is oligomannose N-glycan. The oligomannose N-glycans are increased in high-grade PCa tumors and are associated with more aggressive disease and poorer clinical outcomes [13].

The glycosylation modification impacts the metabolism and immune evasion of tumor cells which affect the PCa development and treatment [15–17]. For instance, the glycosylation modification affects expression of immune checkpoint proteins as the critical targets of cancer immunotherapy [16]. The changes in glycosylation patterns may thus serve as the potential biomarkers for PCa, particularly in predicting disease aggressiveness and response to therapy. Targeting the specific glycosylation pathways represent a novel therapeutic approach for PCa. However, more studies are required to validate these

findings and determine clinical utility.

In this study, the role of glycosylation modification in PCa is explored. The potential of single cell analysis in understanding this complex disease is found out. The implications of glycosylation modification in cancer immunotherapy and development of new treatments for PCa are described. *GALNT7*+ epithelial cells, *UGCG*+ cancer associated macrophages, and other cell subtypes are screened which impact the prognosis and immunotherapies. By gathering the cutting-edge research from multiple disciplines, the cancer biology can be better understood to provide assistance for new advances in PCa diagnosis and treatment.

2. Materials and methods

2.1 Data sources

The scRNA-seq data (GSE141445) for 13 PCa samples was obtained from Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/). The bulk RNA-seq of PRAD for 52 normal and 499 tumor samples were downloaded from UCSC Xena (https://xena.ucsc.edu). The gene set of glycosyltransferases was retrieved from published literature with total of 185 gene [18] (**Supplementary Table 1**). Glycosyltransferases were a class of enzymes transferring sugar molecule from one molecule to another [19]. They were named according to the substrate and type of catalyzed reaction. For example, N-acetylglucosaminyltransferase transferred N-acetylglucosamine (*GlcNAc*) residue from Uridine diphosphate (UDP)-GlcNAc to serine or threonine on protein, forming O-linked glycosylation [20].

2.2 Preprocessing for scRNA-seq

Seurat package (v4.1.1) in R (v4.1, USA) was utilized to preprocess the scRNA-seq data for ensuring downstream analysis quality. Cells with mitochondrial genes of >15% and detected genes of <200 or >4000 were removed. The data was normalized through "LogNormalize" function. The first 2000 highly variable genes were selected for principal component analysis (PCA) and the first 15 PCs for further analysis. The "FindClusters" function (resolution = 0.4) was used to cluster and visualize 15 PCs into Uniform Manifold Approximation and Projection (UMAP). The "FindAllMarkers" function was employed to find the differentially expressed genes in each cluster. Cell type annotation was performed in integrated manner by combining predefined cell markers with online databases including CellMarker (http://xteam.xbio.top/CellMarker/), and PanglaoDB (https://panglaodb.se/).

2.3 Evaluation of metabolism features

ScMetabolism [21] was utilized as an R package to quantify the metabolic activity at single-cell level and to evaluate the metabolic activity of TME in PCa. This package used conventional single-cell matrix files with the VISION algorithm to score each cell, which provided an activity score of each cell in the metabolic pathway. Additionally, the package could calculate the metabolism of each cell type to infer the metabolic activity across entire system.

2.4 Cell-to-cell communication

The "CellChat" (v1.4.0) R package was used to infer the inner relationship of paired ligand-receptor based on the ligand-receptor interaction [22]. CellChat required cell's gene expression data as the input and simulated the probability of cell-to-cell communication by combining gene expression with prior knowledge of interactions between signaling ligands, receptors and their fellow-factors.

2.5 Non-negative Matrix factorization for scRNA-seq

Non-negative Matrix Factorization (NMF) was an unsupervised method decomposing non-negative matrix into the product of two non-negative matrices. NMF algorithm could find non-negative matrices of W and H for any given non-negative matrix V. The non-negative matrix V was almost equal to the product of W and H, *i.e.*, $V \approx W \times H$. The approach identified the underlying patterns and features in the data. NMF clustering was performed for the specific cell type in scRNA-seq data to explore the heterogeneity and functional characteristics within the cells. The "NMF" (v0.24.0) R package was used to conduct the clustering process [23].

2.6 Single-cell transcription factor analysis

Cell heterogeneity in tissues was primarily based on the differences in cell transcriptional states, which were determined and maintained by gene regulatory networks (GRNs) driven by transcription factors. The "SCENIC" (v1.3.1) R package was utilized to calculate the transcription factors activity of scRNA-seq data, which employed the co-expression and motif analysis to compute gene regulatory network reconstruction and identify cell states [24].

2.7 Single sample gene set enrichment analysis

Single sample gene set enrichment analysis (ssGSEA) was a computational method to evaluate the enrichment of set of genes in single sample such as the patient sample. This method was useful for analyzing the gene expression data of small number of samples where traditional methods might not be appropriate because of limited statistical power. In this study, the "FindAllMarkers" function was used to pre-define the differential expression genes of each cluster during cell subtype identification process prior to the NMF clustering. These marker genes were then employed for the ssGSEA analysis of bulk RNA-seq data. "GSVA" R package (v1.46.0) was utilized to perform ssGSEA analysis using "GSVA" function [25].

2.8 Immunotherapy response analysis

Tumor Immune Dysfunction and Exclusion (TIDE; http://tide.dfci.harvard.edu/) was the computational framework designed to evaluate the likelihood of tumor immune escape and rejection based on gene expression profiles of tumor samples [26, 27]. TIDE was utilized in this study for assessing the response to immune checkpoint blocking therapy.

3. Results

3.1 Single-cell RNA-seq data preparation and cell types annotation

The 13 single-cell RNA-seq samples of PCa from GSE141445 were subjected to quality control preprocessing (Fig. 1A). The downstream analysis was performed on 32,603 cells clustered into 22 groups (Fig. 1B). Seven cell types were identified based on this analysis: Myeloids, T cells, Natural killer cells (NK cells), B cells, Epithelial cells, Endothelial cells and Fibroblasts (Fig. 1C). The cell markers for each cell type were designated (Fig. 1D). Transcriptional signatures specific to the cell types were identified. Epithelial cells were enriched for Keratin 19 (*KRT19*), *KRT18* and *KRT8*; T cells and NK cells for CD3 delta subunit of t-cell receptor complex (*CD3D*) and C-X-C motif chemokine receptor 4 (*CXCR4*); Myeloids for Lysozyme (*LYZ*) and *CD68*; Endothelial cells for Transmembrane 4 l six family member 1 (*TM4SF1*) and Von Willebrand Factor (*VWF*); and B cells for *CD69*.

3.2 Cell-to-cell interactions inference

The ligand-receptor interactions were analyzed using CellChat to explore the cell-to-cell interactions (Fig. 2A). CellChat quantified the communications between two cell groups mediated by signaling genes and associated each interaction with the probability value [22]. Significant interactions were identified based on the statistical test that randomly permuted the group labels of cells and recalculated the interaction probability. An intercellular communication network was a weighted directed graph composed of connections between the interacting cell groups. The number of inferred ligand-receptor pairs depended on the method used to calculate average gene expression of cell population. CellChat's default method was "trimean", which produced fewer interactions but helped to find more significant communications. The total number of interactions was higher between epithelial and endothelial cells. The overall strength of interactions was consistent with the results of interaction number, and demonstrated stronger interactions between epithelial and endothelial cells (Fig. 2B). Additionally, the epithelial cells exhibited stronger interactions with myeloids (0.7), NK cells (0.6) and endothelial cells (0.6) compared to the other cells.

The signaling patterns were compared among these cells (Fig. 2C). Epithelial cells had the most outgoing signaling patterns as they sent the highest communication signals to other cells. Myeloids received the largest signaling strength among the incoming signaling patterns. Endothelial cells received the highest signaling pathways compared to the other cells, indicating their involvement in multiple signaling regulations.

3.3 Clustering for the epithelial cells

The epithelial cells were subset and NMF clustering was performed to investigate the heterogeneity of glycosylation in PCa. Fig. 3A showed the identified 10 clusters. The differential gene expression analysis was conducted for each cluster, and results presented in **Supplementary Table 2**. These results were used to identify the glycosylation-related cell subtypes. Cells with highly expressed glycosylation-related genes were confirmed as the glycosylation-positive epithelial cells (Epi). Four glycosylation gene-positive cell types were identified including ALG13 UDP-N-Acetylglucosaminyltransferase Subunit+ (*ALG13*+) Epi-C1, Beta-1,3-Glucuronyltransferase 3+ (*B3GAT3* +) Epi-C2, *UGCG* + Epi-C3 and *GALNT7* + Epi-C5, and one negative cell type, Non-GTs-Epi-C4 (Fig. 3B). Non-GTs-Epi-C4 was the cell population without specific glycosyltransferases as the markers in NMF results.

CellChat was used to evaluate the interactions of epithelial cells subtypes with others. There were more interactions between the epithelial and endothelial cells (Fig. 3C), being consistent with Fig. 2A. The interactions of glycosylation gene-positive epithelial cells with endothelial cells were higher than those with Non-GTs-Epi-C4. Fig. 3D exhibited the interaction strength among these cells which also revealed the higher interaction strength between glycosylation gene-positive epithelial cells and endothelial cells. Additionally, the glycosylation gene-positive epithelial cells displayed higher interaction strengths with other cells.

Resultantly, *ALG13* + Epi-C1, *B3GAT3* + Epi-C2, *UGCG* + Epi-C3 and *GALNT7* + Epi-C5 were identified to interact more with endothelial cells. Glycosylation positive epithelial cells displayed higher metabolic activity than those of the glycosylation negative. *ALG13* (asparagine-linked glycosylation 13) protein was involved in N-linked glycosylation. It worked with *ALG14* to construct functional UDP-GlcNAc glycosyltransferase in endoplasmic reticulum, which catalyzed the protein N-glycosylation second step [28, 29]. Polypeptide-N-acetylgalactosaminyltransferase 7 (*GALNT7*) belonged to GalNAc-transferase family which was previously reported for modifying O-glycosylation, promoting tumor growth in PCa, and transferring GalNAc to Ser/Thr of proteins/peptides [30–32]. UDP-glucose ceramide glycosyltransferase (*UGCG*) was the key enzyme to regulate glycosphingolipid (*GSL*) metabolism and transfer UDP-glucose to ceramide [33].

In conclusion, the glycosylation-related genes could distinguish cell subsets of epithelial cells, and glycosylation-positive epithelial cells had higher communication with other cells, especially the endothelial cells. Metastasis was the major death cause in cancer patients. Tumor cells required to cross endothelial layer of blood vessels to enter blood system and find new growth sites in the body. This process involved the interaction of tumor cells with endothelial cells [34].

3.4 Functional evaluation of epithelial cell subtypes

Further analysis was conducted for the internal regulatory traits of epithelial cell subtypes to clarify their functional features. Given that the epithelial cells had high interactions with endothelial cells, the *VEGF* signaling pathway was focused which promoted the angiogenesis. Glycosylation-positive epithelial cells and fibroblasts were the senders of *VEGF* signaling, while Non-GTs-Epi-C4 did not send this signal (Fig. 4A). Endothelial cells were the only receivers for VEGF pathway, indicating that this signaling pathway had the specialized function in endothelial cells. Furthermore, the specific gene expression of epithelial cell subtypes was investigated,

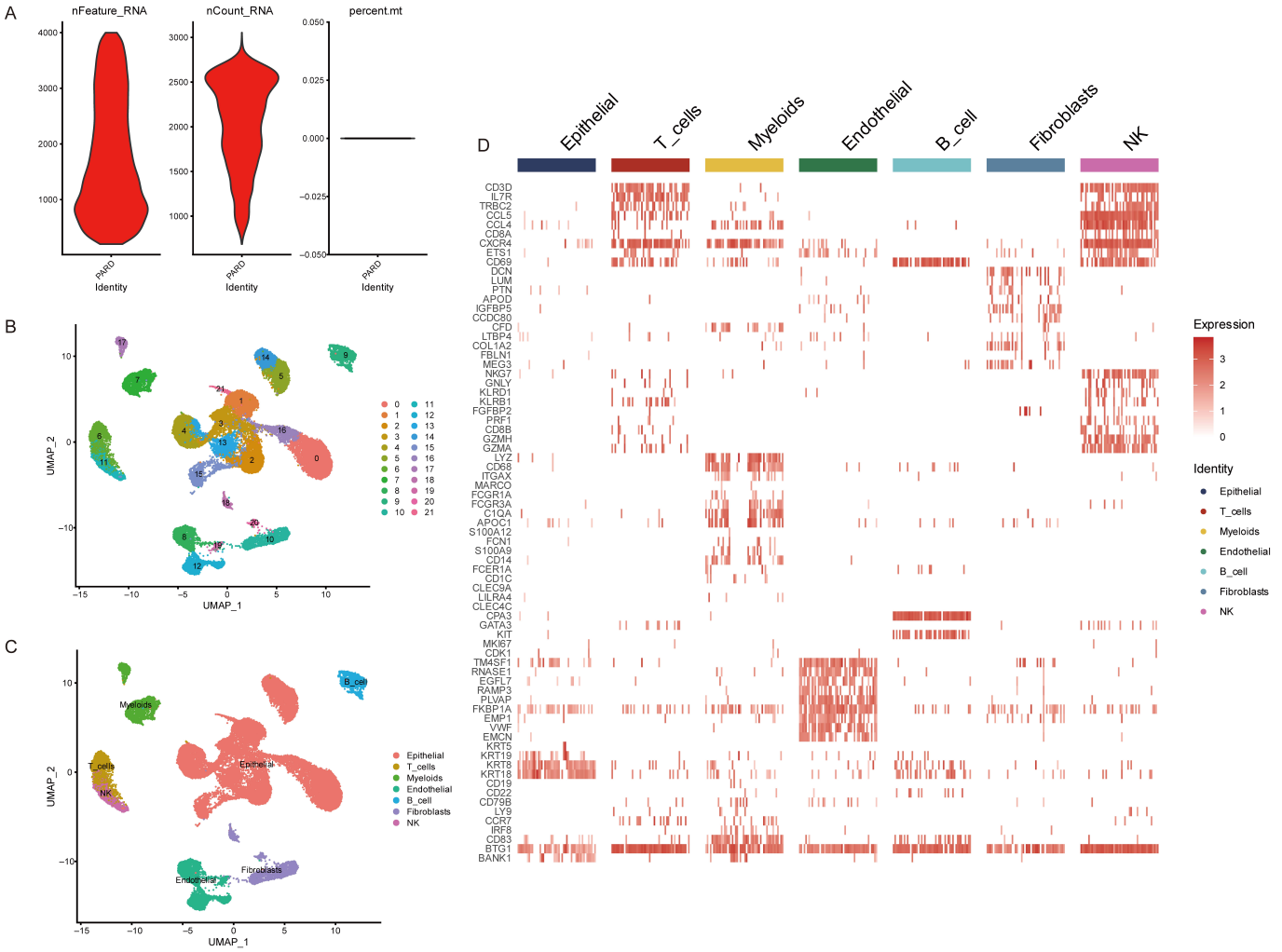

FIGURE 1. Single cell RNA-seq data preparation and cell types annotation. (A) Quality control preprocessing of 13 PCa samples from GSE141445. (B) Clustering analysis of preprocessed data resulting in the identification of 22 groups. (C) Cell type annotation based on the clustering analysis reveal seven distinct cell types: Myeloids, T cells, NK cells, B cells, Epithelial cells, Endothelial cells and Fibroblasts. (D) Cell markers for the identified cell types.

and subsequently projected onto UMAP. Four genes, *ALG13*, *GALNT7*, *B3GAT3* and *UGCG* displayed differences in UMAP distribution to further reflect the intrinsic heterogeneity of epithelial cells in PCa (Fig. 4B).

The metabolism analysis was also performed for epithelial cell subtypes. The metabolic features were calculated for these 5 cell types based on ScMetabolism algorithm. Fig. 4C displayed the metabolic landscape where glycosylation-positive epithelial cells had higher metabolic levels than Non-GTs-Epi-C4. Furthermore, the glycosylation-positive epithelial cells exhibited heterogeneity with the others. *ALG13* + Epi-C1, and *B3GAT3* + Epi-C2 manifested the highest metabolic levels followed by *UGCG* + Epi-C3 and *GALNT7* + Epi-C5.

The glycosylation-positive epithelial cells thus had higher communication with endothelial cells in *VEGF* signaling, but not the Non-GTs-Epi-C4. Additionally, the glycosylation-positive epithelial cells exhibited high metabolic status features.

3.5 Clustering for cancer associated endothelial cells (CAE)

Endothelial cells displayed high interactions with epithelial cells. They were subset and NMF clustering was performed based on the glycosylation genes expression profiles. Ten clusters were identified as shown in Fig. 5A. The differential gene expression analysis was conducted on each cluster (**Supplementary Table 3**). Resultantly, 6 glycosylation gene-positive endothelial cell subtypes were identified (Fig. 5B), including Alpha-1,3-Mannosyl-Glycoprotein 2-Beta-N-Acetylglucosaminyltransferase (*MGAT1*) + CAE-C1, UGCG + CAE-C3, ST6 Beta-Galactoside Alpha-2,6-Sialyltransferase 1 (*ST6GAL1*) + CAE-C4, ALG5 + CAE-C5, Beta-1,4-Galactosyltransferase 1 (*B4GALT1*) + CAE-C6 and Chondroitin Sulfate Synthase 1 (*CHSY1*) + CAE-C7.

CellChat calculated the interaction profiles between endothelial cell subtypes and epithelial cells. All the glycosylation gene-positive endothelial cells displayed higher number of interactions with epithelial cells compared to Non-GTs-CAE-C2 (Fig. 5C). *MGAT1* + CAE-C1, *UGCG* + CAE-C3, *ALG5* + CAE-C5 and *B4GALT1* + CAE-C6 exhibited

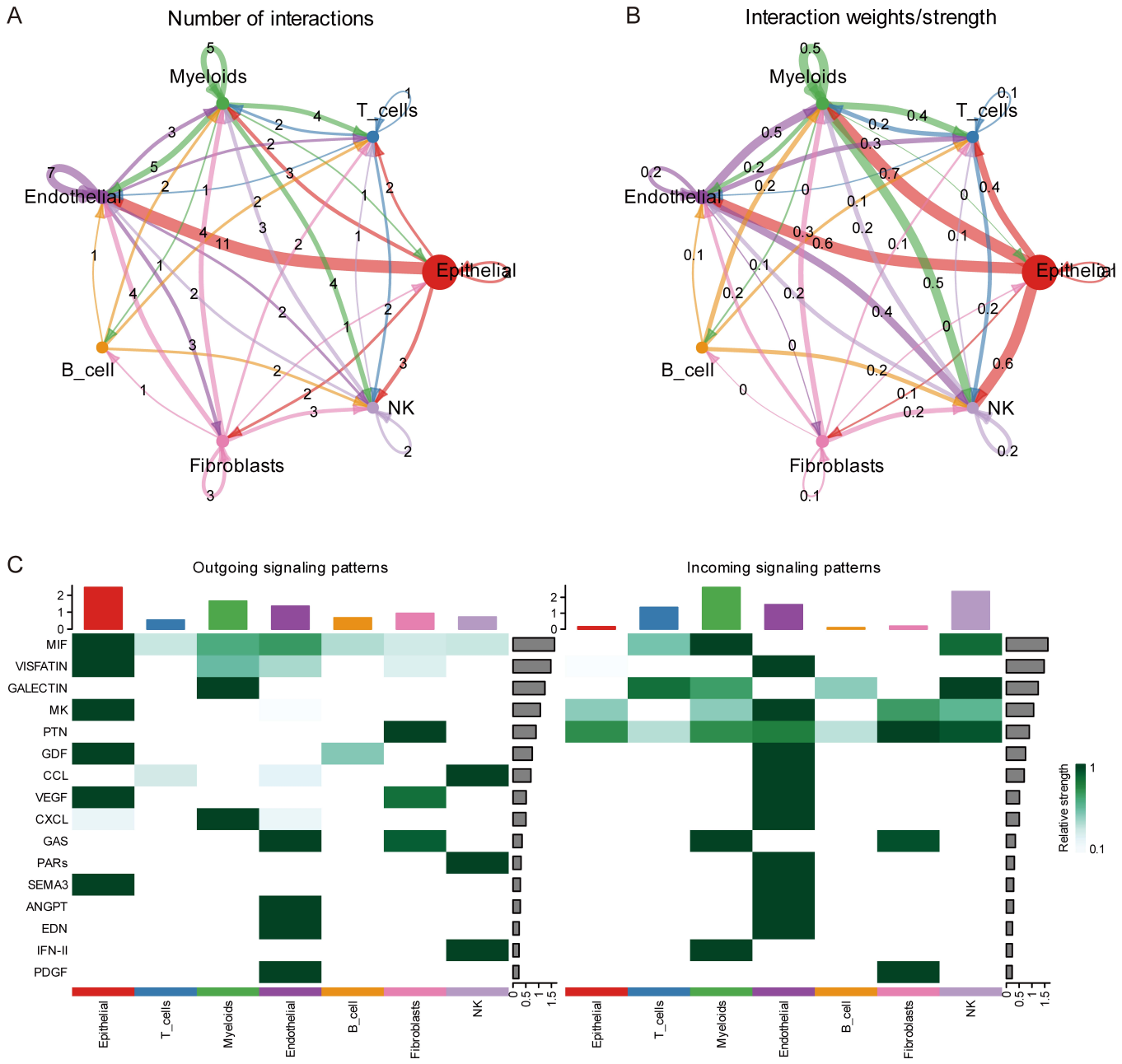

FIGURE 2. Cell-to-cell interactions inference. (A) Analysis of ligand-receptor paired interactions using CellChat, revealing the cell-to-cell interaction numbers. (B) Overall strength between various cell types. (C) Signaling patterns among different cell types. *MIF*: Macrophage Migration Inhibitory Factor; *MK*: Midkine; *PTN*: Pleiotrophin; *GDF*: Growth Differentiation Factor; *CCL*: C-C Motif Chemokine Ligand; *VEGF*: Vascular Endothelial Growth Factor; *CXCL*: C-X-C Motif Chemokine Ligand; *GAS*: Gastrin; *PARs*: Par-3 Family Cell Polarity Regulator; *SEMA3*: Semaphorin 3E; *ANGPT*: Angiopoietin; *EDN*: Endothelin; *IFN-II*: Interferon-II; *PDGF*: Platelet Derived Growth Factor.

more interactions than other subtypes. The interaction strength with epithelial cells was also higher in glycosylation-positive endothelial cells than in Non-GTs-CAE-C2 (Fig. 5D). The findings suggested that glycosylation-positive endothelial cells might promote epithelial cell function through their interactions.

It was found that VEGF signaling pathway had an impact on endothelial cells by the epithelial cells (Fig. 4A). The affected specific endothelial cell subtypes were investigated as illustrated in Fig. 5E. The findings indicated that epithelial cells mainly influenced *ALG5* + CAE-C5, *B4GALT1* + CAE-C6, *CHSY1* + CAE-C7, *MGAT1* + CAE-C1 and *UGCG* + CAE-C3 endothelial cell subtypes through VEGF signaling pathway. These glycosylation-positive subtypes exhibited high interaction levels with epithelial cells in VEGF signaling pathway, implying their role in promoting epithelial cell function.

The transcriptional activity of these subtypes was evaluated by the Scenic analysis. A higher status of transcriptional activity was found in the glycosylation-positive endothelial cells than in Non-GTs-CAE-C2 (Fig. 5F). These findings suggested that glycosylation might regulate endothelial cell function and serve as potential therapeutic target for PCa.

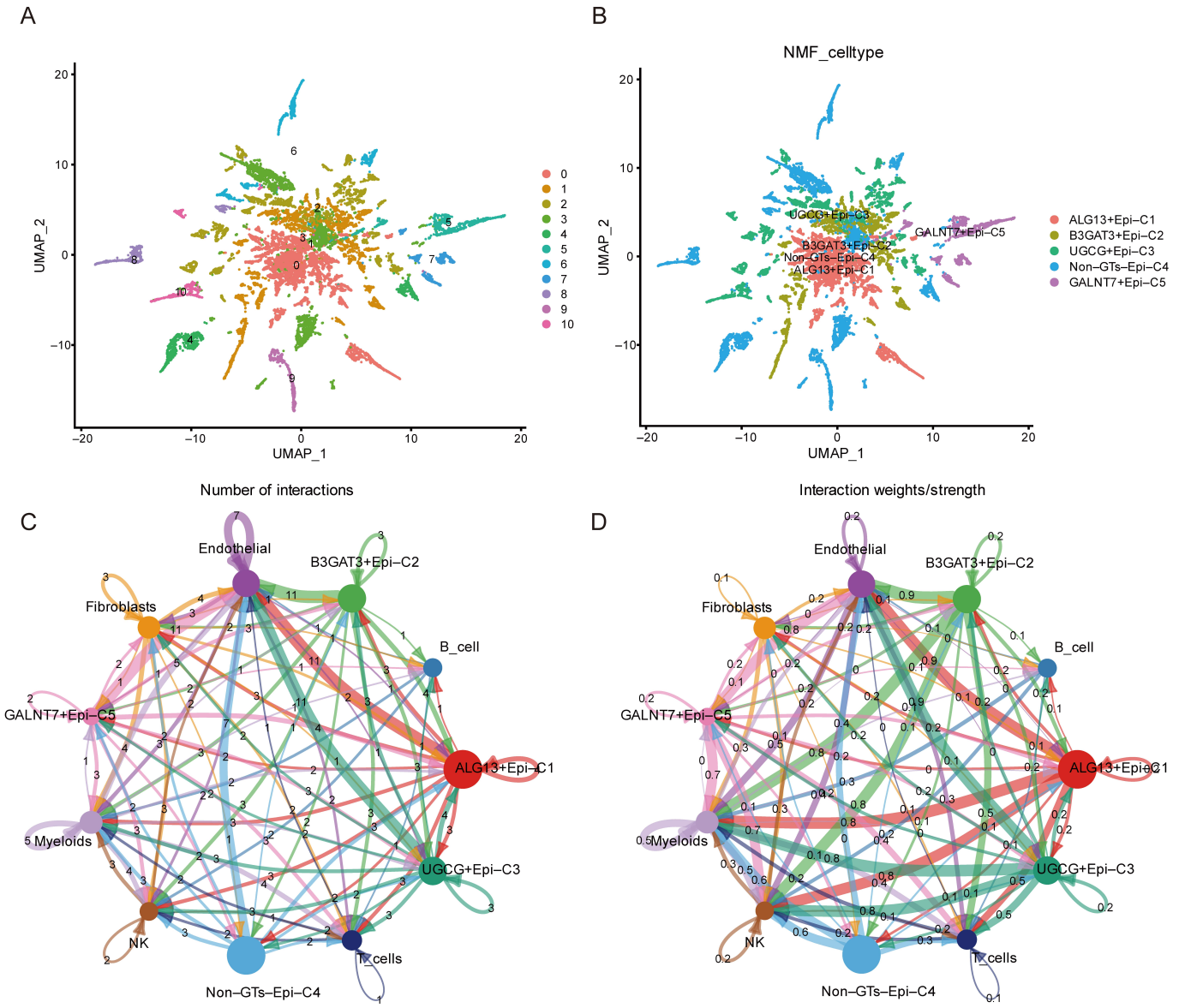

FIGURE 3. NMF clustering for the epithelial cells. (A) NMF clustering of epithelial cells resulting in the identification of 10 clusters. (B) Identification of glycosylation gene-positive cell types in epithelial clusters including *ALG13* + Epi-C1, *B3GAT3* + Epi-C2, *UGCG* + Epi-C3, *GALNT7* + Epi-C5, and one negative cell type, Non-GTs-Epi-C4. (C) Evaluation of cell-to-cell interactions using CellChat, showing more interactions between epithelial cells and endothelial cells. (D) Interaction strength among different cell types.

3.6 Clustering macrophages and their interactions with epithelial cells

This study identified that myeloid and NK cells had role in the TME of PCa through high interactions with epithelial cells. Macrophages had complex phenotypes and represented the most abundant innate immune population in TME [35]. Cancer-associated macrophages (CAM) regulated the tumor growth, metastasis and drug resistance. Myeloid cells were thus reannotated as illustrated in Fig. 6A. Two cell types, *i.e.*, macrophages and dendritic cells were identified.

NMF clustering was performed on macrophages based on their glycosylation gene expression. Differential gene expression analysis was conducted with each cluster (**Supplementary Table 4**). Two cell types, *MGAT4A* + CAM-C1 and *UGCG* + CAM-C3 were identified. CellChat was used to assess their interactions with epithelial cells. Higher number of interactions and interaction strength were observed between the glycosylation-positive macrophages and epithelial cells (Fig. 6B,C). The metabolic characteristics of each macrophage subtype were also evaluated. *MGAT4A* + CAM-C1 and *UGCG* + CAM-C3 exhibited higher metabolism than Non-GTs-CAM-C2 (Fig. 6D). These findings suggested that various glycosylation macrophage subtypes had distinct metabolic profiles, which might influence their functions in cell-to-cell interactions and metabolism.

3.7 Scoring cell subtypes by ssGSEA in bulk RNA-seq

The marker genes of each cell subtype (**Supplementary Table 5**) were integrated to perform ssGSEA based on the results

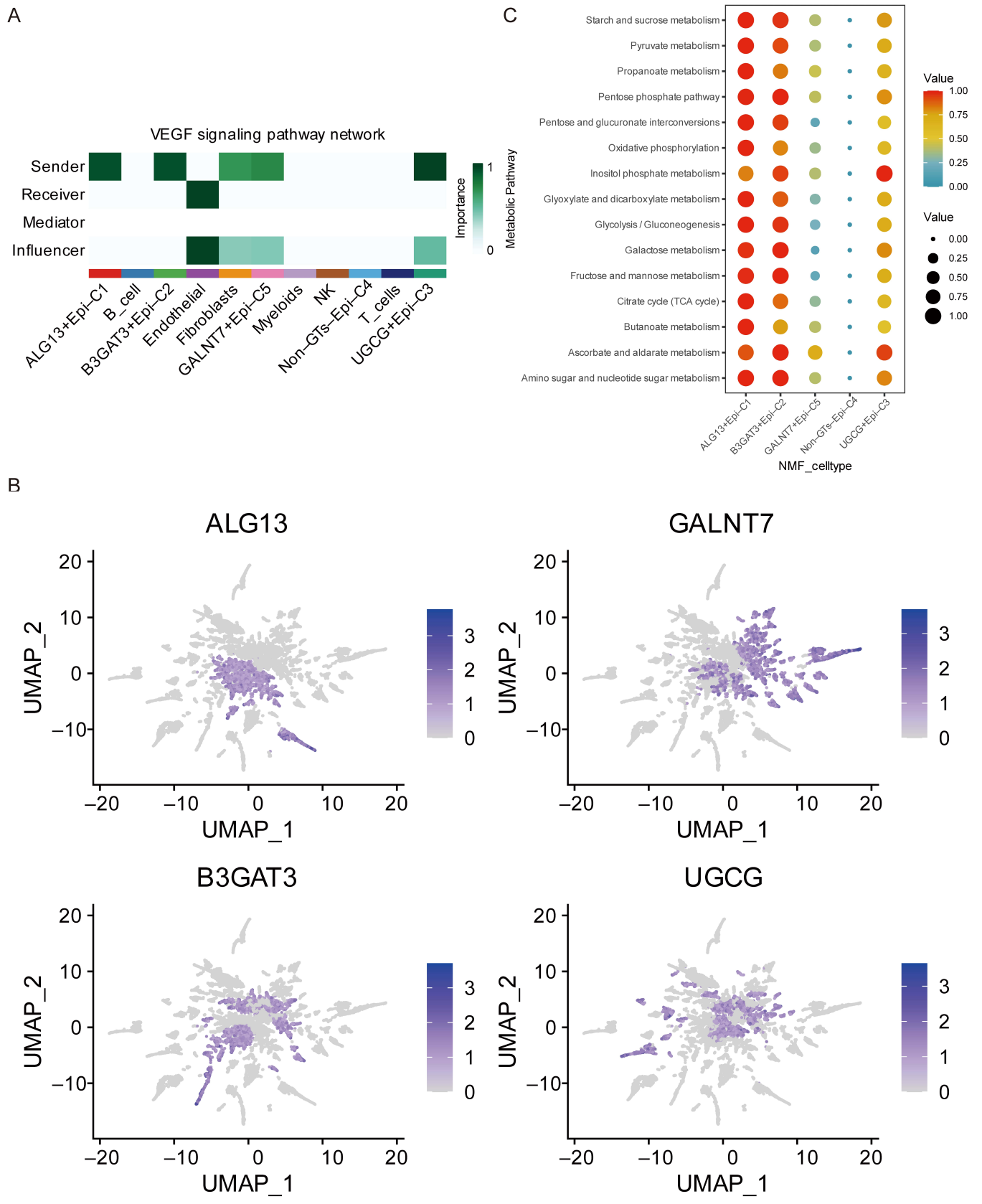

FIGURE 4. Functional evaluation of epithelial cell subtypes. (A) Evaluation of the *VEGF* signaling pathway. (B) Projection of specific gene expressions of epithelial cell subtypes (*ALG13*, *GALNT7*, *B3GAT3*, *UGCG*) onto UMAP, indicating the intrinsic heterogeneity of epithelial cells in PCa. (C) Metabolic landscape analysis of the epithelial cell subtypes. *GALNT7*: Polypeptide N-Acetylgalactosaminyltransferase 7.

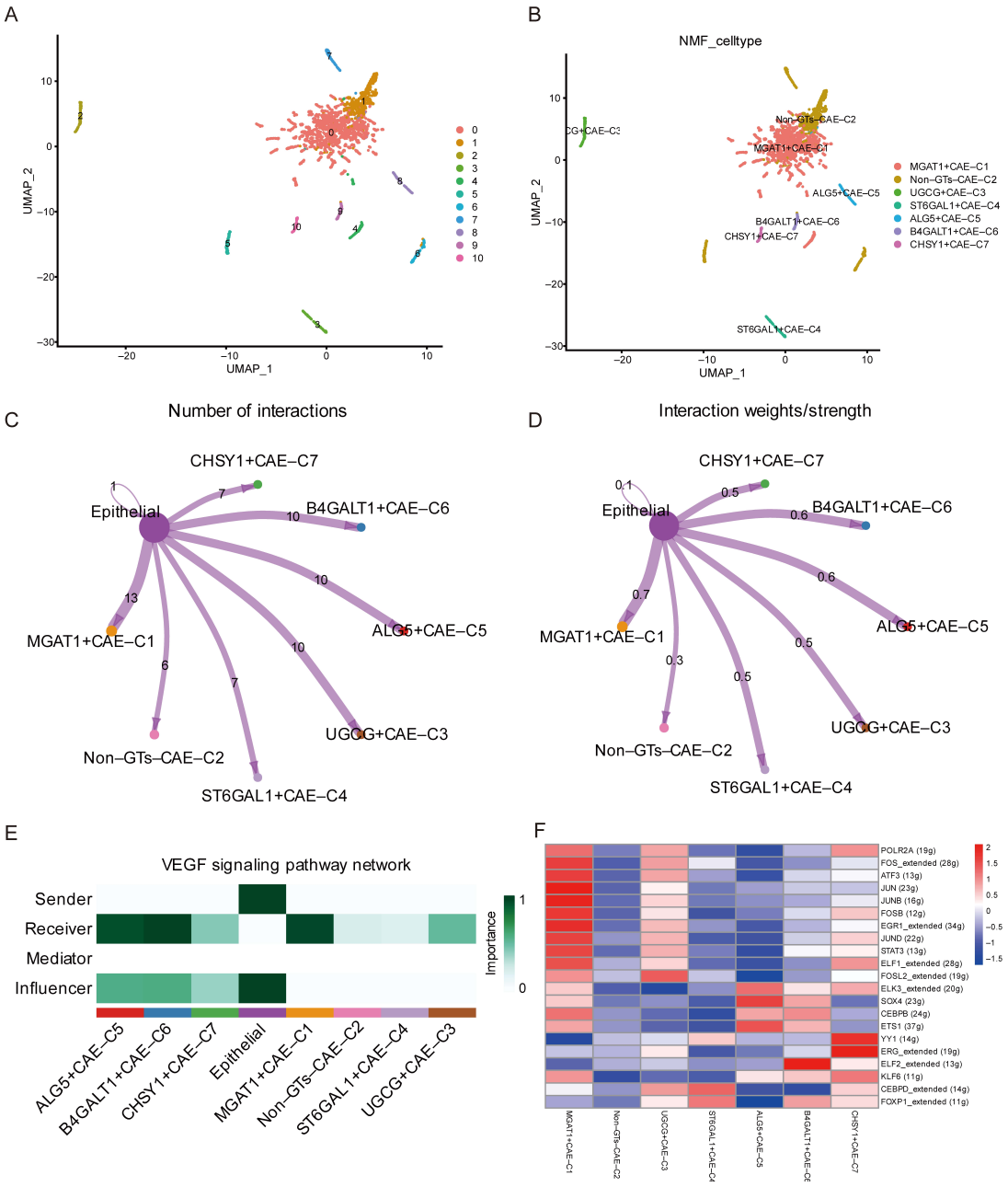

FIGURE 5. Clustering and interactions of endothelial cells with epithelial cells in PCa. (A) NMF clustering of endothelial cells identified 10 clusters based on the glycosylation gene expression profiles. (B) Differential gene expression analysis reveal 6 glycosylation gene-positive endothelial cell subtypes. (C) Interaction profiles between glycosylation gene-positive endothelial cell subtypes and epithelial cells, showing higher number of interactions compared to Non-GTs-CAE-C2. (D) Interaction strength with epithelial cells being higher in glycosylation-positive endothelial cells than in Non-GTs-CAE-C2. (E) Epithelial cells influenced the specific glycosylation-positive endothelial cell subtypes through *VEGF* signaling pathway. (F) Higher transcriptional activity observed in glycosylation-positive endothelial cells compared to Non-GTs-CAE-C2. *POLR2A*: RNA Polymerase II Subunit A; *FOS*: Fos Proto-Oncogene, *AP*-1 Transcription Factor Subunit; *ATF3*: Activating Transcription Factor 3; *JUN*: Jun Proto-Oncogene; *JUNB*: JunB Proto-Oncogene; *FOSB*: FosB Proto-Oncogene; *EGR1*: Early Growth Response 1; *JUND*: JunD Proto-Oncogene; *STAT3*: Signal Transducer And Activator Of Transcription 3; *ELF1*: E74 Like ETS Transcription Factor 1; *FOSL2*: FOS Like 2; *ELK3*: ETS Transcription Factor ELK3; *FOSL2*: FOS Like 2; *ELK3*: ETS Transcription Factor ELK3; *SOX4*: SRY-Box Transcription Factor 4; *CEBPB*: CCAAT Enhancer Binding Protein Beta; *ETS1*: ETS Proto-Oncogene 1; *YY1*: YY1 Transcription Factor; *ERG*: ETS-related gene; *ELF2*: E74 Like ETS Transcription Factor 2; *KLF6*: KLF Transcription Factor 6; *CEBPD*: CCAAT Enhancer Binding Protein Delta; *FOXP1*: Forkhead Box P1.

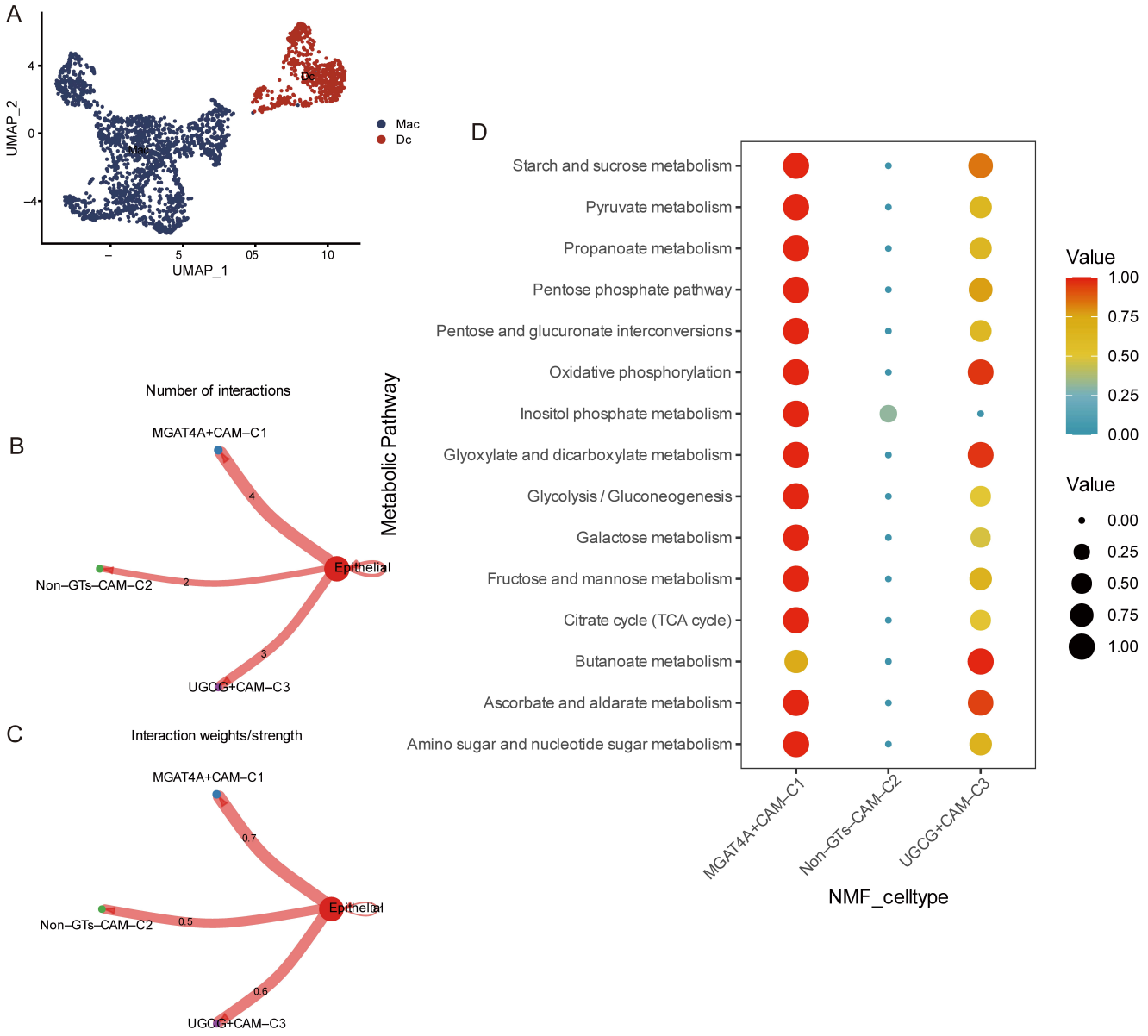

FIGURE 6. Clustering and interactions of macrophages with epithelial cells in PCa. (A) NMF clustering of macrophages identified two cell types: macrophages and dendritic cells. (B) Interaction profiles between glycosylation-positive macrophages and epithelial cells, showing higher number of interactions compared to Non-GTs-CAM-C2. (C) Interaction strength with epithelial cells being higher in glycosylation-positive macrophages than in Non-GTs-CAM-C2. (D) Metabolic characteristics of macrophage subtypes with *MGAT4A* + CAM-C1 and *UGCG* + CAM-C3 exhibiting higher metabolism than Non-GTs-CAM-C2.

of differential gene expression analysis. Analysis results reflected the relative cell scoring in bulk RNA-seq data. The cell scoring of *UGCG* + CAM-C3, *ALG13* + Epi-C1, *B3GAT3* + Epi-C2 and *GALNT7* + Epi-C5 were higher in prostate tumor than in normal samples (Fig. 7A). The scoring of *UGCG* + Epi-C3, *B4GALT1* + CAE-C6 and *CHSY1* + CAE-C7 were lower in tumor compared to the normal samples (p value of < 0.05 was statistically significant, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$ and ****$p < 0.0001$).

The immunotherapy response analysis was performed on bulk RNA-seq data using TIDE. The efficacy of ICB treatment response in PCa tumor samples was compared. Higher *UGCG* + CAM-C3, *B3GAT3* + Epi-C2 and *GALNT7* + Epi-C5 scores were associated with greater efficacy of ICB treatment, while higher *ST6GAL1* + CAE-C4 and ALG5 + CAE-C5 scores were linked to the lesser response (Fig. 7B). The efficacy of ICB therapy might thus be influenced by specific glycosylation-related cell subtypes present in PCa.

3.8 Prognosis efficacy of glycosylation-related cell subtypes

This study provided evidence that glycosylation-related cell subtypes might regulate PCa therapy and affect patient prognosis and survival. A survival analysis was conducted using glycosylation-related cell subtypes. The results showed that high *ALG13* + Epi-C1 and *GALNT7* + Epi-C5 scores in

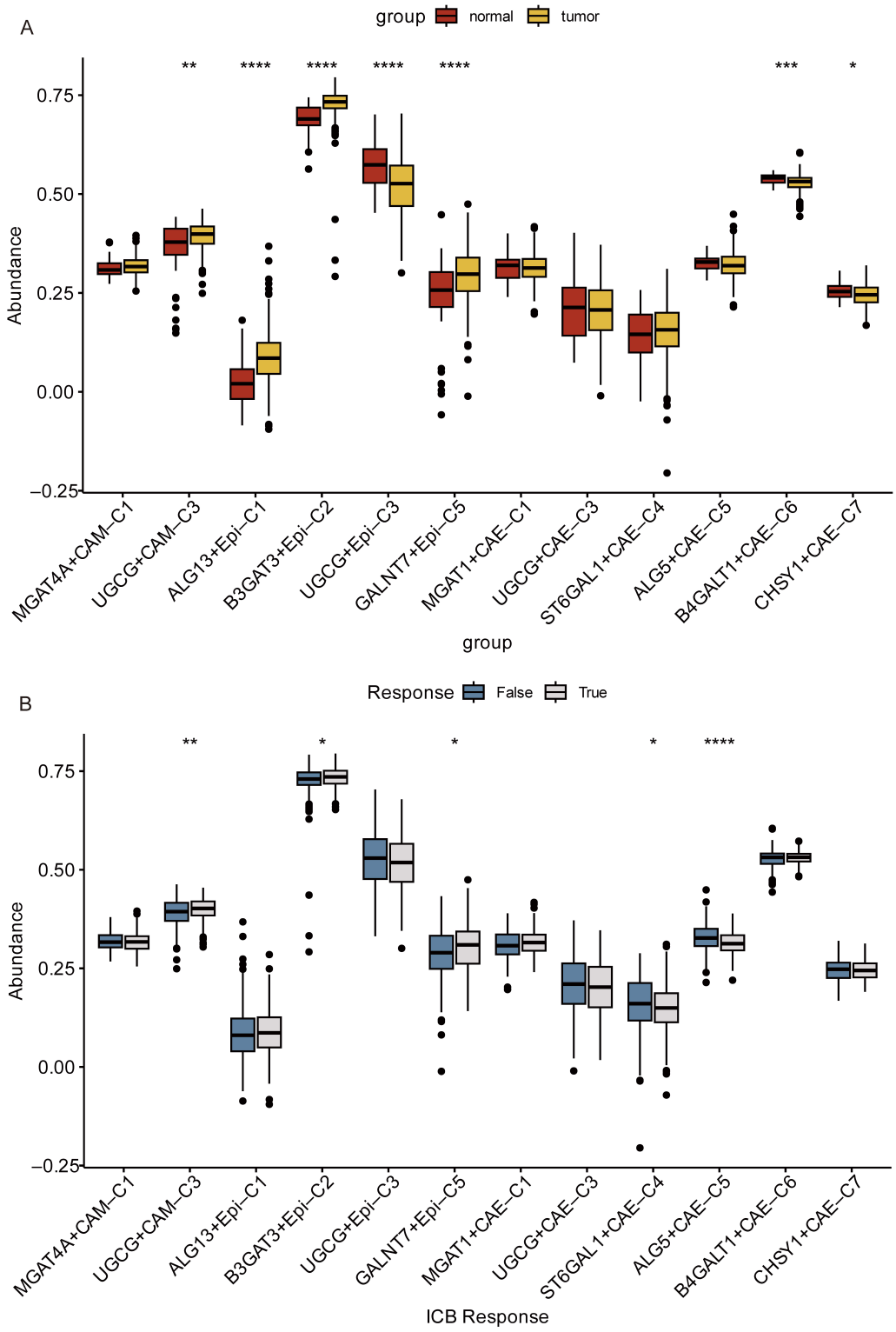

FIGURE 7. Cell subtype scoring and immunotherapy response in PCa. (A) Cell scoring of cell subtypes in prostate tumor compared to the normal samples. (B) Immunotherapy response analysis showing association between cell subtype scores and ICB treatment efficacy.

epithelial cell subtypes were associated with poor prognosis (Fig. 8A,B). An impact on the prognosis from endothelial cell subtypes was also observed. Higher *B4GALT1* + CAE-C6 scores were associated with adverse effects, while *ALG5* + CAE-C5 and *UGCG* + CAE-C3 scores were linked to the improved prognosis (Fig. 8C–E). Furthermore, the macrophages also had influence on prognosis where high *UGCG* + CAM-C3 scores contributed to poor prognosis (Fig. 8F). The glycosylation-related cell subtypes had thus an impact on prognosis of PCa patients. This study identified differences in the effects of glycosylation-related cell subtypes among various cell types.

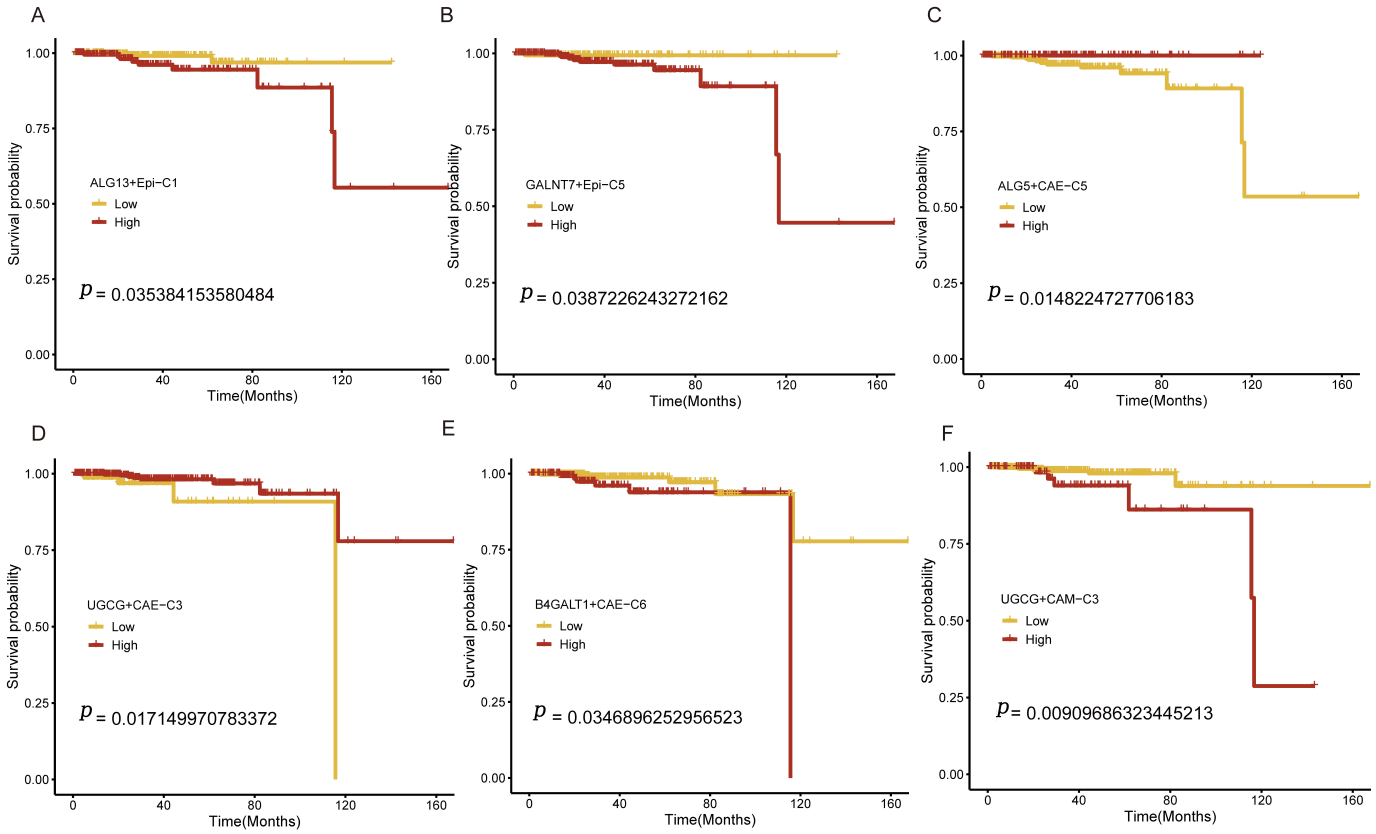

FIGURE 8. Prognosis efficacy of glycosylation-related cell subtypes in PCa. (A,B) High *ALG13* + Epi-C1 and *GALNT7* + Epi-C5 scores in epithelial cell subtypes associated with poor prognosis. (C–E) High *UGCG* + CAE-C3 and *ALG5* + CAE-C5 scores in endothelial cell subtypes associated with improved prognosis while high *B4GALT1* + CAE-C6 scores in endothelial cell subtypes associated with adverse effects on prognosis. (F) High *UGCG* + CAM-C3 scores in macrophages associated with poor prognosis.

4. Discussion

ScRNA-seq had provided new perspective at the single-cell level. A better understanding of inner traits was attained based on the scRNA-seq data analysis of PCa patients. In this study, scRNA-seq and bulk RNA-seq were utilized to focus on the influence of glycosylation in dissecting cellular components' inner heterogeneity of PCa.

Seven cell types were identified in the scRNA-seq of PCa. Further analysis of cell-to-cell interactions within TME revealed that the epithelial cells had high level of interaction with endothelial cells. NMF clustering of the epithelial cells was performed based on the glycosylation genes expression to investigate whether intrinsic tumor cells exhibited intrinsic heterogeneity. The survival analysis depicted that *ALG13* + Epi-C1 and *GALNT7* + Epi-C5 were linked with the poor prognosis of PCs patients. *ALG13* regulated $GABA_A$ receptors and was related to epileptogenesis [36]. *ALG13* gene mutation caused early infantile epileptic encephalopathy known as *ALG13*-CDG, which was a congenital disorder of glycosylation (CDG) resulting from disorder in N-linked protein glycosylation [37]. However, there was limited work on the *ALG13* role in tumors. A neuroblastoma study had shown that ALG13 was associated with poor clinical outcomes, however its specific mechanism was unclear [38]. In this study, it was identified that high ALG13 expression in epithelial cells of prostate cancer (PCa) was associated with poor clinical outcomes. Similarly, it was found that PCa patients with *GALNT7* positive epithelial cells had poor prognosis. Targeted knockdown of *GALNT7* could inhibit the proliferation, migration and invasion in nasopharyngeal carcinoma [39]. Therefore, it was proposed that *GALNT7* was a potential target for PCa treatment.

Endothelial cells had role in the TME of PCa because of high interactions with epithelial cells. Therefore, NMF clustering of endothelial cells was performed and interaction analysis was conducted with the epithelial cells. Results showed an overall upregulation of communication with epithelial cells in glycosylation-positive endothelial subtypes. It was established based on the ligand-receptor interaction analysis that *VEGF* signaling pathway had role in the process of epithelial cells signaling to endothelial cells. The *VEGF* signaling pathway was a key factor of promoting angiogenesis in cancer [40], which enhanced the tumor growth and evaded detection [41]. The findings suggested that endothelial cells had role in the TME of PCa *via* the angiogenesis regulation. A study on PCa treatment showed that downregulating *VEGF*-related genes promoted the apoptosis in prostate cancer cell lines [42]. A glycosylation-related endothelial cells profile was established at the single-cell level. The high transcriptional activity in glycosylation-positive endothelial subtypes suggested that glycosylation enhanced inner function of endothelial cells in the TME. The survival analysis results indicated that gene-

specific endothelial cells had varying influences. Endothelial cells positive for *ALG5* and *UGCG* exhibited protective effect on prognosis, while those positive for *B4GALT1* did not. These differences might ascribe to their varying functions. *B4GALT1*, a beta-1,4-galactosyltransferase and newly identified glycosyltransferase of *PD-L1*, was an enzyme of the glycosylation process [43]. High *B4GALT1* expression in pancreatic cancer was associated with chemotherapy resistance and cancer progression [43]. This enzyme expression could be regulated by p65 activation, which in turn interacted with *CDK11* p110 protein *via* N-linked glycosylation to promote cancer progression and chemoresistance [44].

The outcomes similar to those in epithelial and endothelial cells were observed based on the NMF and interaction analysis results on macrophages. Macrophages tested positive for glycosylation exhibited high interactions with epithelial cells, which were also characterized by the high metabolic levels. However, *UGCG*-positive macrophages had negative impact on survival prognosis, whereas they were the protective factor in endothelial cells. In breast cancer, *UGCG* regulated the glutamine metabolism and oxidative phosphorylation to promote tumor proliferation and drug resistance [45, 46]. *UGCG* was also reported to be associated with lysosomal autophagy and identified as a potential cancer target [33]. The survival differences between macrophages and endothelial cells caused by *UGCG* reflected that this gene might play various roles which required further elucidations. *UGCG* was involved in the endothelial cells, epithelial cells and macrophages NMF clustering, demonstrating its multiple roles and critical impact in PCa immune microenvironment.

Immunotherapy had treated a variety of cancers including PCa. However, not all the patients responded equally to immunotherapy. It was imperative to identify biomarkers for predicting treatment response. In this study, TIDE analysis was used to investigate the efficacy of immune checkpoint blockade (ICB) treatment in PCa patients based on glycosylation-related cell subtypes. The results depicted that glycosylation could alter the ICB efficacy. Different effects of various genes were observed, indicating that glycosylation modifications performed diverse functions. Further studies were required on these genes to impart more benefits to PCa patients.

Overall, these findings highlighted the complex and dynamic nature of TME in PCa. Understanding the interactions between different cell types was crucial in developing effective therapies. Further research might elucidate the mechanisms underlying these interactions and their potential clinical implications. Some biomarkers were also found which might identify the patients less likely to benefit from ICB treatment and thus require alternative treatments. These findings could lead to more personalized treatment strategies for PCa patients with improved outcomes.

5. Conclusions

This study provides insights to the glycosylation role in prostate cancer subtyping. The identification of various cell biomarkers and subtypes based on the glycosylation patterns opens up avenues for targeted therapy and personalized treatment of this disease, such as *GALNT7*+ epithelial cells and *UGCG*+ cancer associated macrophages. These findings have implications in the development of therapeutic strategies with improved patient outcomes and reduced PCa disease burden worldwide. This work may inspire further investigations into the role of glycosylation for other cancers and develop effective therapies.

AVAILABILITY OF DATA AND MATERIALS

The data presented in this study are available on reasonable request from the corresponding author.

AUTHOR CONTRIBUTIONS

GXD and WHZ—designed the research study. SJT, WHZ and JZ—performed the research. SJT—analyzed the data. SJT and GXD—wrote the manuscript. All authors contributed to editorial changes in the manuscript. All authors read and approved the final manuscript.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

ACKNOWLEDGMENT

Not applicable.

FUNDING

This research was supported by Shanghai Municipal Health Commission under Grant No. 202240376.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

SUPPLEMENTARY MATERIAL

Supplementary material associated with this article can be found, in the online version, at https://oss.jomh.org/files/article/1773592329684172800/attachment/Supplementary%20material.xlsx.

REFERENCES

[1] Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. CA: A Cancer Journal for Clinicians. 2022; 72: 7–33.

[2] Sandhu S, Moore CM, Chiong E, Beltran H, Bristow RG, Williams SG. Prostate cancer. The Lancet. 2021; 398: 1075–1090.

[3] Hansen AR, Massard C, Ott PA, Haas NB, Lopez JS, Ejadi S, *et al*. Pembrolizumab for advanced prostate adenocarcinoma: findings of the KEYNOTE-028 study. Annals of Oncology. 2018; 29: 1807–1813.

[4] Antonarakis ES, Piulats JM, Gross-Goupil M, Goh J, Ojamaa K, Hoimes CJ, *et al*. Pembrolizumab for treatment-refractory metastatic castration-resistant prostate cancer: multicohort, open-label phase II KEYNOTE-199 study. Journal of Clinical Oncology. 2020; 38: 395–405.

[5] Smith BAH, Bertozzi CR. The clinical impact of glycobiology: targeting selectins, siglecs and mammalian glycans. Nature Reviews Drug Discovery. 2021; 20: 217–243.

[6] Berger M, Kaup M, Blanchard V. Protein glycosylation and its impact on biotechnology. Genomics and Systems Biology of Mammalian Cell Culture. 2011; 28: 165–185.

[7] Copoiu L, Malhotra S. The current structural glycome landscape and emerging technologies. Current Opinion in Structural Biology. 2020; 62: 132–139.

[8] Trbojević-Akmačić I, Lageveen-Kammeijer GSM, Heijs B, Petrović T, Deriš H, Wuhrer M, *et al*. High-throughput glycomic methods. Chemical Reviews. 2022; 122: 15865–15913.

[9] Pinho SS, Reis CA. Glycosylation in cancer: mechanisms and clinical implications. Nature Reviews Cancer. 2015; 15: 540–555.

[10] Mehboob MZ, Lang M. Structure, function, and pathology of protein O-glucosyltransferases. Cell Death & Disease. 2021; 12: 71.

[11] Reily C, Stewart TJ, Renfrow MB, Novak J. Glycosylation in health and disease. Nature Reviews Nephrology. 2019; 15: 346–366.

[12] Kaluza A, Szczykutowicz J, Ferens-Sieczkowska M. Glycosylation: rising potential for prostate cancer evaluation. Cancers. 2021; 13: 3726.

[13] Gilgunn S, Murphy K, Stöckmann H, Conroy PJ, Murphy TB, Watson RW, *et al*. Glycosylation in indolent, significant and aggressive prostate cancer by automated high-throughput n-glycan profiling. International Journal of Molecular Sciences. 2020; 21: 9233.

[14] Scott E, Munkley J. Glycans as biomarkers in prostate cancer. International Journal of Molecular Sciences. 2019; 20: 1389.

[15] Xu Y, Gao Z, Hu R, Wang Y, Wang Y, Su Z, *et al*. PD-L2 glycosylation promotes immune evasion and predicts anti-EGFR efficacy. Journal for ImmunoTherapy of Cancer. 2021; 9: e002699.

[16] Shi C, Wang Y, Wu M, Chen Y, Liu F, Shen Z, *et al*. Promoting anti-tumor immunity by targeting TMUB1 to modulate PD-L1 polyubiquitination and glycosylation. Nature Communications. 2022; 13: 6951.

[17] Adeva-Andany M, Pérez-Felpete N, Fernández-Fernández C, Donapetry-García C, Pazos-García C. Liver glucose metabolism in humans. Bioscience Reports. 2016; 36: e00416.

[18] Li Y, Lin Y, Aye L, Dong L, Zhang C, Chen F, *et al*. An integrative pan-cancer analysis of the molecular and biological features of glycosyltransferases. Clinical and Translational Medicine. 2022; 12: e872.

[19] Lairson LL, Henrissat B, Davies GJ, Withers SG. Glycosyltransferases: structures, functions, and mechanisms. Annual Review of Biochemistry. 2008; 77: 521–555.

[20] Saha A, Bello D, Fernandez-Tejada A. Advances in chemical probing of protein O-GlcNAc glycosylation: structural role and molecular mechanisms. Chemical Society Reviews. 2021; 50: 10451–10485.

[21] Wu Y, Yang S, Ma J, Chen Z, Song G, Rao D, *et al*. Spatiotemporal immune landscape of colorectal cancer liver metastasis at single-cell level. Cancer Discovery. 2022; 12: 134–153.

[22] Jin S, Guerrero-Juarez CF, Zhang L, Chang I, Ramos R, Kuan CH, *et al*. Inference and analysis of cell-cell communication using CellChat. Nature Communications. 2021; 12: 1088.

[23] Gaujoux R, Seoighe C. A flexible R package for nonnegative matrix factorization. BMC Bioinformatics. 2010; 11: 367.

[24] Aibar S, González-Blas CB, Moerman T, Huynh-Thu VA, Imrichova H, Hulselmans G, *et al*. SCENIC: single-cell regulatory network inference and clustering. Nature Methods. 2017; 14: 1083–1086.

[25] Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-Seq data. BMC Bioinformatics. 2013; 14: 7.

[26] Fu J, Li K, Zhang W, Wan C, Zhang J, Jiang P, *et al*. Large-scale public data reuse to model immunotherapy response and resistance. Genome Medicine. 2020; 12: 21.

[27] Jiang P, Gu S, Pan D, Fu J, Sahu A, Hu X, *et al*. Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. Nature Medicine. 2018; 24: 1550–1558.

[28] Gao X, Moriyama S, Miura N, Dean N, Nishimura S. Interaction between the C termini of Alg13 and Alg14 mediates formation of the active UDP-N-acetylglucosamine transferase complex. Journal of Biological Chemistry. 2008; 283: 32534–32541.

[29] Averbeck N, Gao X, Nishimura S, Dean N. Alg13p, the catalytic subunit of the endoplasmic reticulum UDP-GlcNAc glycosyltransferase, is a target for proteasomal degradation. Molecular Biology of the Cell. 2008; 19: 2169–2178.

[30] Scott E, Hodgson K, Calle B, Turner H, Cheung K, Bermudez A, *et al*. Upregulation of GALNT7 in prostate cancer modifies O-glycosylation and promotes tumour growth. Oncogene. 2023; 42: 926–937.

[31] Masone MC. The role of GALNT7 as a potential diagnostic marker in prostate cancer. Nature Reviews Urology. 2023; 20: 198.

[32] Schjoldager KT, Narimatsu Y, Joshi HJ, Clausen H. Global view of human protein glycosylation pathways and functions. Nature Reviews Molecular Cell Biology. 2020; 21: 729–749.

[33] Jain V, Harper SL, Versace AM, Fingerman D, Brown GS, Bhardwaj M, *et al*. Targeting UGCG overcomes resistance to lysosomal autophagy inhibition. Cancer Discovery. 2023; 13: 454–473.

[34] Bolik J, Krause F, Stevanovic M, Gandraß M, Thomsen I, Schacht S, *et al*. Inhibition of ADAM17 impairs endothelial cell necroptosis and blocks metastasis. Journal of Experimental Medicine. 2022; 219: e20201039.

[35] Christofides A, Strauss L, Yeo A, Cao C, Charest A, Boussiotis VA. The complex role of tumor-infiltrating macrophages. Nature Immunology. 2022; 23: 1148–1156.

[36] Huo J, Ren S, Gao P, Wan D, Rong S, Li X, *et al*. ALG13 participates in epileptogenesis *via* regulation of $GABA_A$ receptors in mouse models. Cell Death Discovery. 2020; 6: 87.

[37] Ng BG, Eklund EA, Shiryaev SA, Dong YY, Abbott MA, Asteggiano C, *et al*. Predominant and novel de novo variants in 29 individuals with ALG13 deficiency: clinical description, biomarker status, biochemical analysis, and treatment suggestions. Journal of Inherited Metabolic Disease. 2020; 43: 1333–1348.

[38] De Antonellis P, Carotenuto M, Vandenbussche J, De Vita G, Ferrucci V, Medaglia C, *et al*. Early targets of miR-34a in neuroblastoma. Molecular & Cellular Proteomics. 2014; 13: 2114–2131.

[39] Nie G, Luo L, Duan H, Li X, Yin M, Li Z, *et al*. GALNT7, a target of miR-494, participates in the oncogenesis of nasopharyngeal carcinoma. Tumor Biology. 2016; 37: 4559–4567.

[40] Carmeliet P. VEGF as a key mediator of angiogenesis in cancer. Oncology. 2005; 69: 4–10.

[41] Hicklin DJ, Ellis LM. Role of the vascular endothelial growth factor pathway in tumor growth and angiogenesis. Journal of Clinical Oncology. 2005; 23: 1011–1027.

[42] Mirzaei A, Rashedi S, Akbari MR, Khatami F, Aghamir SMK. Combined anticancer effects of simvastatin and arsenic trioxide on prostate cancer cell lines *via* downregulation of the VEGF and OPN isoforms genes. Journal of Cellular and Molecular Medicine. 2022; 26: 2728–2740.

[43] Zhang J, Zhang G, Zhang W, Bai L, Wang L, Li T, *et al*. Loss of RBMS1 promotes anti-tumor immunity through enabling PD-L1 checkpoint blockade in triple-negative breast cancer. Cell Death & Differentiation. 2022; 29: 2247–2261.

[44] Chen Y, Su L, Huang C, Wu S, Qiu X, Zhao X, *et al*. Galactosyltransferase B4GALT1 confers chemoresistance in pancreatic ductal adenocarcinomas by upregulating N-linked glycosylation of $CDK11^{p110}$. Cancer Letters. 2021; 500: 228–243.

[45] Schömel N, Hancock SE, Gruber L, Olzomer EM, Byrne FL, Shah D, *et al*. UGCG influences glutamine metabolism of breast cancer cells. Scientific Reports. 2019; 9: 15665.

[46] Schömel N, Gruber L, Alexopoulos SJ, Trautmann S, Olzomer EM, Byrne FL, *et al*. UGCG overexpression leads to increased glycolysis and increased oxidative phosphorylation of breast cancer cells. Scientific Reports. 2020; 10: 8182.

How to cite this article: Shijun Tong, Wenhui Zhu, Jing Zhai, Guanxiong Ding. Exploring the significance of glycosylation in prostate cancer subtyping through single-cell analysis. Journal of Men's Health. 2024; 20(3): 32-44. doi: 10.22514/jomh.2024.036.