

FULL LENGTH ARTICLE

Uncovering genes driving developmental stage progression in prostate cancer through spatial transcriptomics

Yongjun Quan^a, Mingdong Wang^a, Fan Zou^a, Hong Zhang^b,
Yishan Zhang^a, Yongchen Jin^a, Hao Ping^{a,*}

^a Department of Urology, Beijing Tongren Hospital, Capital Medical University, Beijing 100176, China

^b Department of Pathology, Beijing Tongren Hospital, Capital Medical University, Beijing 100176, China

Received 22 February 2025; accepted 11 November 2025

Available online 11 December 2025

KEYWORDS

Glandular epithelial (GE) cells;
Gleason score (GS);
H2AFJ;
Prostate cancer (PCa);
SLC4A4;
Spatial transcriptomics (ST)

Abstract Precisely delineating the transcriptomic profiles of glandular epithelial (GE) cells in prostate cancer (PCa) remains a significant challenge primarily due to their diffuse and multi-focal distribution. To address this, we employed spatial transcriptomics (ST) to analyze 12 PCa tissue samples from 10 patients, aiming to identify PCa progression-associated genes by analyzing expression patterns across histologically distinct regions. Transcriptomic classification via principal component analysis (PCA), uniform manifold approximation and projection (UMAP), and Louvain clustering revealed spatially resolved histological structures within each tissue section. The malignancy status, progression stages, and developmental trajectories of GE clusters were further assessed using inferred copy number variation (inferCNV), diffusion pseudotime (DPT), and partition-based graph abstraction (PAGA) analyses. Based on the preliminary characterization of developmental trajectories, pairwise comparisons of GE clusters identified key oncogenes—including TFF3, OR51E2 (PSGR), FOLH1 (PSMA), AMACR (P504S), FOS (a subunit of AP-1), SLC4A4, EGR1, NDUFB9, and H2AFJ—that are positively associated with PCa progression. Immunohistochemistry (IHC) validation further confirmed the elevated expression of SLC4A4 and H2AFJ in advanced-stage PCa. Overall, this study establishes an ST-based framework for predicting PCa progression and provides valuable insight for the identification of progression-associated genes holding promise as clinical biomarkers.

© 2025 The Authors. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

* Corresponding author.

E-mail address: drpinghao@126.com (H. Ping).

Peer review under the responsibility of Chongqing Medical University.

Introduction

In 2025, prostate cancer (PCa) is estimated to account for 313,780 new cases and 35,770 deaths in the United States, establishing it as the most prevalent cancer and the second leading cause of cancer-related mortality among men.¹ These statistics underscore the substantial public health burden posed by PCa and emphasize the urgent need for further research to elucidate its underlying molecular mechanisms and to advance the development of more effective therapeutic strategies.

PCa is widely recognized as a multifocal disease that typically originates within a single gland.^{2,3} Glandular epithelial (GE) cells, including basal, luminal, and neuroendocrine cells, have been identified as the origin of PCa.^{4–11} Within prostate tissue, PCa frequently exhibits a diffuse distribution of multiple foci, which are often interspersed with non-glandular components, such as the fibromuscular stroma. The intratumor heterogeneity observed in multifocal PCa is hypothesized to arise from either a shared ancestor clone^{12–16} or independent multiclonal origins.^{17–20} Each lesion may represent a distinct stage in the evolutionary trajectory of PCa. Consequently, analyzing genetic patterns across multiple lesions is essential for elucidating the origins of PCa and identifying key drivers of its progression.

Accurately delineating the gene expression profiles of GE cells in PCa remains challenging due to the poorly defined macroscopic boundaries between tumor and adjacent non-tumor tissues. Although single-cell sequencing (scRNA-seq) enables transcriptomic profiling at the cellular level,^{21–23} this approach is limited by its reliance on computational inference and the lack of spatial context, which is essential for resolving the complex histological architecture of PCa and its surrounding microenvironment.^{24,25}

Spatial transcriptomics (ST) overcomes these limitations by integrating RNA sequencing with spatial localization, thereby allowing direct correlation of transcriptomic profiles with specific histological features and tissue morphology.^{16,26–32} The Visium ST platform achieves high spatial resolution through microarrays with 4992 spatially barcoded spots (each 55 μm in diameter) distributed over a $6.5 \times 6.5 \text{ mm}^2$ capture area. The Visium CytAssist ST is an advanced technology that further enhances accuracy in gene expression mapping by employing precise tissue transfer and multiple specific probe pairs to capture high-quality target sequences with reduced background noise.^{27,30} This improvement enables robust integration of spatial transcriptomic data with distinct pathological structures, thereby providing deeper insights into spatial gene activity and underlying biological processes.^{16,29,31–36}

Prior to analyzing multifocal PCa, it is essential to conduct a standardized evaluation of the progression stage for each lesion based on hematoxylin-eosin (HE)-stained images. The Gleason grading system remains the most reliable approach for assessing the pathological stage and progression of PCa.^{37–40} This system enables the prediction of PCa malignancy and offers critical prognostic insights, including the risk of extraprostatic extension, biochemical recurrence, and overall survival.^{41–46} The Gleason score (GS), which is derived from the primary and secondary

Gleason patterns, ranges from $1 + 1$ to $5 + 5$. PCa patients with a GS sum of 7 or higher are classified as having aggressive disease, which is associated with a high risk of extraprostatic extension and an unfavorable prognosis.^{47,48}

In our previous research, we performed preliminary Visium ST analysis on a limited number of PCa tissue samples to identify genes potentially associated with PCa progression.^{29,31,32} These candidate genes were subsequently validated through bioinformatic databases, histopathological assessments, and cellular functional experiments.^{29,31,32} However, the restricted sample size in these earlier investigations may have limited the ability to fully capture inter-individual heterogeneity, potentially introducing inaccuracies into the findings. In the present study, we expanded the scope of analysis to include 12 PCa tissue samples obtained from 10 patients, employing both Visium and Visium CytAssist ST platforms to investigate transcriptome expression patterns across PCa tissues exhibiting diverse GE clusters. Key genes linked to GS progression were identified and further validated through immunohistochemistry (IHC) in an independent cohort of 90 PCa tissue samples.

Materials and methods

Collection of prostate cancer (PCa) tissues

A cohort of ninety PCa patients who had undergone prostatectomy was recruited for this study. Their clinicopathological data are detailed in Table 1. Both PCa and prostatic intraepithelial neoplasia (PIN) tissues from these patients were extracted and stored into the 10% neutral formalin, and then embedded in paraffin for subsequent IHC analysis.

Eight PCa samples extracted from six patients were subjected to Visium CytAssist spatial transcriptomics (ST) analysis (Table 2). Additionally, Visium ST data from four PCa samples in our previous study²⁹ were reanalyzed and integrated into the current dataset. These tissue extraction sites were precisely identified using prostate magnetic resonance imaging (MRI) and confirmed through pathological diagnosis of prostate needle core biopsies. In two cases, bilateral tumors with differing GSs were observed, raising the possibility of independent tumor formation in each lobe; consequently, two samples were collected from each of these tissues. Fresh tissue specimens were snap-frozen on dry ice, embedded in optimal cutting temperature (OCT) compound, and stored at -80°C until further analysis.

All the human prostate tissue samples were obtained from Beijing Tongren Hospital. Informed consent was obtained from all patients, and the study was approved by the Ethics Committee of Beijing Tongren Hospital, affiliated with Capital Medical University (Approval No. TREC2023-KY030).

Visium CytAssist spatial transcriptomics (ST) (10x genomics visium)

Tissue sections embedded in OCT with an RNA integrity number (RIN) ≥ 4 were cryosectioned at a thickness of 10 μm and mounted on Superfrost slides (VWR, Radnor, Pennsylvania, USA) according to Demonstrated Protocol CG000636. The spatial microarray on the slide contained 4992 distinct barcoded spots within capture areas measuring

6.5 × 6.5 mm². Each spot had a diameter of 55 µm and a center-to-center distance of 100 µm, providing transcriptome information representing the gene expression within the tissue (Fig. 1A). Tissue sections mounted on the slides were fixed in pre-cooled methanol solution at −20 °C for 30 min, followed by HE staining and microphotography using the Olympus BX53 microscope (Olympus, Shinjuku-ku, Tokyo, Japan) and were subsequently destained as per Demonstrated Protocol CG000614. Probe hybridization, wash, ligation, and release procedures were then performed on the destained tissue slides. Next, the slides with tissue sections were processed using the Visium CytAssist instrument (10x Genomics, CA, USA), during which the probes were transferred to the capture area of the Visium CytAssist spatial gene expression slide. Probe extension, pre-amplification, and library construction were completed following the Visium CytAssist Spatial Gene Expression Reagent Kits User Guide (CG000495). Sequencing was conducted on the Illumina NovaSeq X Plus platform, with sequencing libraries demultiplexed with bcl2fastq (Illumina, CA, USA). The resulting FASTQ files were aligned to the human reference genome (GRCh38) using Space Ranger v3.0.

Analysis of gene expression using scanpy package in python

The transcriptomic profiling data generated by Space Ranger were analyzed using the Python (version 3.12) package Scanpy. In the data filtering phase, a maximum threshold for mitochondrial gene proportion was assigned to each sample based on its highest observed value (rounded to two decimal places), a strategy designed to incorporate virtually all spatial spots with detectable gene expression. The read count data were normalized to ensure consistent total expression across all spatial spots and subsequently log-transformed using the natural logarithm [$\ln(1 + \text{normalized value})$] to improve the distribution normality and mitigate the impact of highly expressed genes. Highly variable genes (HVGs) were identified using the following thresholds: a mean expression range of 0.0125–3 and a dispersion range of 0.5–1000. Additionally, an image processing algorithm was applied to the captured region of the PCa tissue section based on spatial barcode information from the pre-processed data and the sequencing reads corresponding to each spatial spot.

Dimensionality reduction

To reduce the dimensionality of the transcriptomic data from PCa sections, we performed principal component analysis (PCA) using HVGs as input and retained the top 50 principal components for subsequent analysis. The detailed methodology is detailed in the Scikit-learn documentation (<https://scikit-learn.org/stable/modules/generated/sklearn.decomposition.PCA.html>). Leveraging the PCA-reduced data, we then applied uniform manifold approximation and projection (UMAP) to visualize the global transcriptomic structure in two-dimensional space. Following this, Louvain clustering analysis was employed to categorize the data points into classification ns, allowing for the

Table 1 Clinicopathological characteristics of prostate cancer patients for IHC analysis.

Clinicopathological parameters	Total (n = 90)
Age	
Median age, year (IQR)	68.5 (65–73.25)
Range, year (Min, Max)	54–83
Number of age category, n (%)	
<65	19 (21.1)
≥65	71 (78.9)
Pre-RP t-PSA level	
Median, ng/mL (IQR)	9.65 (5.66–21.375)
Range, ng/mL (Min, Max)	0.1–160
Number of t-PSA status, n (%)	
<4 ng/mL	8 (8.9)
4–10 ng/mL	38 (42.2)
10–20 ng/mL	20 (22.2)
>20 ng/mL	24 (26.7)
f-PSA/t-PSA (f/t)	
Median (IQR)	0.1185 (0.1–0.17)
Range (Min, Max)	0.0–0.5
Number of f/t status, n (%)	
<0.16	63 (70.0)
≥0.16	27 (30.0)
Number of total GS status, n (%)	
<7	20 (22.2)
=7	36 (40.0)
>7	34 (37.8)
Number of pT-stage, n (%)	
T2	53 (58.9)
T3a	18 (20.0)
T3b	11 (12.2)
T4	8 (8.9)
Number of ISUP grade, n (%)	
1	20 (22.2)
2	20 (22.2)
3	16 (17.8)
4	19 (21.1)
5	15 (16.7)
Number of AJCC stage, n (%)	
I	9 (10.0)
IIA	7 (7.8)
IIB	16 (17.8)
IIC	8 (8.9)
IIIA	11 (12.2)
IIIB	24 (26.7)
IIIC	15 (16.7)

IQR: Interquartile range; RP, radical prostatectomy; t-PSA: total prostate-specific antigen; f-PSA: free prostate-specific antigen; GS: Gleason score; pT: pathological T; ISUP: International Association of Urology Pathology; AJCC: American Joint Committee on Cancer.

simultaneous alignment of transcriptomic patterns with histological structures.

Inferred copy number variation (inferCNV) analysis

The acquisition of specific genomic aberrations has been linked to the progression of PCa to advanced stages.^{12–16,49,50} CNV analysis is instrumental in elucidating

clonal hierarchies and assessing tumor malignancy.^{12–16,22,23} In this study, somatic large-scale chromosomal CNVs were inferred by employing transcriptomic data obtained from ST analysis, facilitated through a Python reimplementation of inferCNV (<https://infercnvpy.readthedocs.io/en/latest/infercnv.html>). The expression intensity of genes across positions in the whole genome was examined in relation to a designated set of “normal” reference spots. We utilized PIN as the primary reference group. In instances where PIN cells were unavailable, cells with a GS of 3 + 3 were used as an alternative reference group due to their well-differentiated characteristics and low metastatic potential,^{51–53} which have been proposed to represent pre-cancerous lesions.^{54,55} In cases where neither PIN nor GS (3 + 3) were available, the group with the lowest GS was selected as the reference group. The inferCNV values of local genomic fragments across 22 autosomes were visualized in a heatmap, while the median CNV signals within each cluster were depicted in boxplots. This approach was designed to elucidate the cellular origins^{12–16} and malignancy-associated features of each cluster.^{22,23}

Diffusion pseudotime (DPT) analysis

DPT enables the inference of cellular progression in biological processes from snapshot data by leveraging geodesic distance along the graph structures.⁵⁶ This approach allows for the reconstruction of the developmental progression of PCa cells, facilitating the identification of transient or metastable states, branching decisions and differentiation endpoints.⁵⁶ It is implemented within the Scanpy framework (<https://scanpy.readthedocs.io/en/stable/generated/scanpy.tl.dpt.html>).⁵⁷

Partition-based graph abstraction (PAGA) analysis

PAGA is a graph abstraction that reconciles clustering with trajectory inference through a topology preserving map of single cells.⁵⁸ PAGA generates an interpretable, graph-like map (PAGA graph) of the arising data manifold, based on estimating connectivity of manifold partitions (<https://github.com/theislab/paga>).

PAGA maps preserve the global topology of data, enable multi-resolution analysis, and significantly enhance the computational efficiency of the typical exploratory data analysis workflow. It not only captures both discrete disconnected and continuous connected cell-to-cell variations but also reliably predicts developmental trajectories and gene expression dynamics using transcriptomic datasets. The algorithms used to generate PAGA were previously described in a prior study⁵⁸ and have been implemented within the Scanpy framework (<https://scanpy.readthedocs.io/en/stable/generated/scanpy.tl.paga.html>).

Differentially expressed gene (DEG) analysis

DEG analysis was conducted to thoroughly explore the heterogeneity of gene expression across various developmental stages of histological clusters in PCa. DEGs between the specified clusters were identified using the Wilcoxon test implemented in `scanpy.tl.rank_genes_groups` (https://scanpy.readthedocs.io/en/stable/generated/scanpy.tl.rank_genes_groups.html). For this analysis, filtering thresholds were set such that the absolute value of the natural logarithm of the fold change ($|\ln FC|$) > 0.25, and the false discovery rate (FDR) as determined by the Benjamini–Hochberg procedure (Q-value) < 0.0001. The distribution of DEGs across various clusters was summarized and visualized using volcano plots and Venn diagrams.

Score a set of genes

The composite score for a specific gene set was calculated using the `scanpy.tl.score_genes` function, as outlined in the Scanpy documentation (https://scanpy.readthedocs.io/en/stable/generated/scanpy.tl.score_genes.html). The score is computed as the difference between the mean expression of a target gene set and that of a reference gene set. The reference set is generated by random sampling from the `gene_pool` based on binned expression values. This methodology, initially developed in Seurat,²⁴ was implemented in Scanpy by Davide Cittaro. The gene score was

Table 2 Clinicopathological characteristics of selected PCa patients for ST analysis.

No.	Age	TPSA (ng/mL)	FPSA/TPSA	GS (pathology)	pT	ISUP grade	AJCC stage
Cytassist_1	68	16.8	0.089	3 + 4	T2	2	IIB
Cytassist_2	66	4.5	0.111	4 + 3	T2	3	IIC
Cytassist_3	65	27.1	0.192	4 + 3	T3b	3	IIIB
Cytassist_4							
Cytassist_5	65	40.9	0.071	4 + 3	T4	3	IIIB
Cytassist_6	69	12.2	0.066	4 + 5	T3a	5	IIIC
Cytassist_7							
Cytassist_8	66	10.6	0.1	3 + 4	T2	2	IIB
Visium_1	65	23.91	0.1	3 + 4	T3b	2	IIIB
Visium_2	71	32.4	0.259	4 + 4	T3b	4	IIIB
Visium_3	63	58.88	0.086	5 + 4	T2	5	IIIC
Visium_4	72	9.86	0.175	4 + 3	T2	3	IIC

TPSA: total prostate specific antigen; FPSA: free prostate specific antigen; GS: Gleason score; ISUP: International Association of Urology Pathology; AJCC: American Joint Committee on Cancer.

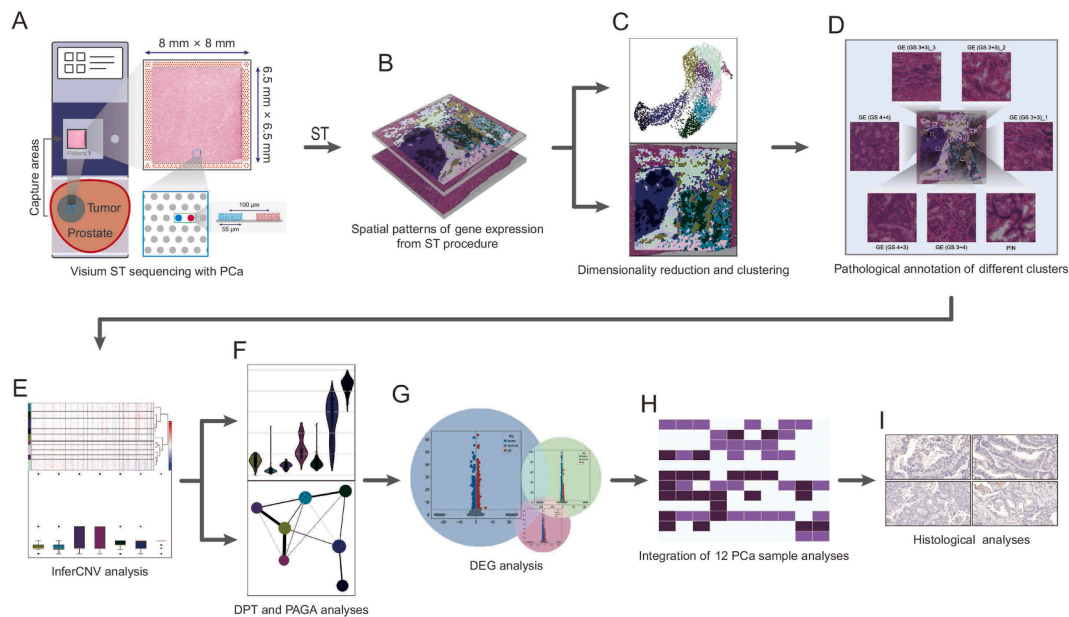


Figure 1 Workflow of this study. (A) Cryosectioned prostate cancer (PCa) tissue samples were mounted on Superfrost slides (VWR) containing a spatial microarray of 4992 barcoded spots (55 μm in diameter, 100 μm center-to-center spacing) to enable spatially resolved gene expression profiling. (B) Transcriptomic read count matrices were generated following the CytAssist ST platform. (C) Dimensionality reduction and clustering of the spatial spots were performed through PCA, UMAP and Louvain clustering analyses. (D) The clusters were annotated as distinct histological structures based on pathological evaluation. (E) Malignancy-associated features of the classified clusters were assessed through inferCNV analysis. (F) Developmental trajectories and progression stages were further characterized using DPT and PAGA analyses. (G) To identify genes associated with PCa progression, DEG analysis was conducted by comparing GE clusters aligned with GS progression patterns that met all predefined criteria. (H) Commonly dysregulated DEGs across 12 PCa samples were identified and summarized. (I) The protein expression of candidate DEGs (H2AFJ, SLC4A4, TFF3) was validated in PCa tissues spanning a spectrum of GSs and PIN using IHC staining.

utilized to comprehensively evaluate the expression patterns of the selected gene sets across various stages of histological structure.

Enrichment analysis

The biological functions, signaling pathways, and diseases associated with the identified DEGs were elucidated through Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses. These analyses were conducted using R software (version 4.3.1) with the “clusterProfiler”, “enrichplot”, “org.Hs.eg.db”, and “ggplot2” packages.

Bioinformatics analysis of the cancer genome Atlas prostate adenocarcinoma (TCGA-PRAD) dataset

The input gene set was validated using data from the TCGA-PRAD cohort. Clinicopathological characteristics and follow-up information were obtained from two primary sources: the Genomic Data Commons (GDC) Data Portal (<https://portal.gdc.cancer.gov/>) and the UCSC Xena platform (<https://xenabrowser.net/datapages/>).⁵⁹ These comprehensive datasets provided robust phenotypic and clinical annotations essential for subsequent analyses. Gene expression read counts from TCGA-PRAD patients were normalized to transcripts per kilobase million (TPM) to account for variations in sequencing depth and gene length. Subsequently, key clinicopathological variables, including

GS, pathological T stage (pT), and TP53 mutation status, were statistically analyzed using R software (version 4.3.1) with the “limma”, “reshape2”, and “ggpubr” packages. Kaplan–Meier survival analysis and Cox proportional hazards regression were performed to evaluate two prognostic outcomes: recurrence-free survival (RFS) and progression-free survival (PFS). The corresponding RFS and PFS data were extracted from the phenotype files, as previously detailed.^{59,60} Survival analyses were conducted using R software (version 4.3.1) with the “survival” and “survminer” packages.

Immunohistochemistry (IHC) analysis

IHC analysis was performed to confirm the protein expression of the representative oncogenes. Paraffin-embedded tissue samples were sectioned into 5 μm slices and subjected to antigen retrieval. The sections were subsequently incubated with specific primary antibodies, followed by the appropriate secondary antibodies. Detailed information on the antibodies used is provided in Table S1.

Statistical analysis

Statistical analyses were conducted using Python (version 3.12), R software (version 4.3.1), SPSS (version 23; IBM, Armonk, NY, USA), GraphPad Prism (version 9.0.0; GraphPad Software, La Jolla, CA, USA), and Microsoft Excel 2019 (Microsoft Corp., Redmond, WA, USA). Continuous variables

were compared between two groups using the Wilcoxon test or Student's *t*-test, along with *P*- or *Q*-value calculations. For comparisons among three or more groups, the Kruskal–Wallis test or one-way/two-way analysis of variance (ANOVA) was applied, followed by *post hoc* tests, including Tukey's test or Dunn's multiple comparison test. Survival analysis was conducted using Kaplan–Meier curves with the log-rank test and univariate Cox regression models to calculate hazard ratios (HRs). *P*-values or *Q*-values less than 0.05 were considered statistically significant.

To quantify the strength of the IHC staining and perform the statistical analysis, representative images were selected at 36.4× magnification. The staining intensity was scored on a scale of 1–3 (1: weak, 2: moderate, 3: strong), and the percentage of positively stained cells was scored as follows: 0 for ≤ 5%, 1 for 6%–25%, 2 for 26%–50%, 3 for 51%–75%, and 4 for ≥ 76%. The staining index (SI) was then calculated as the product of the staining intensity score and the staining percentage score, following previously established methods.^{61,62}

Results

Workflow of the Visium CytAssist ST analysis

Figure 1 illustrates the workflow of the spatial transcriptomics analysis in this study. The extracted PCa tissues were cryosectioned and mounted on Superfrost slides (VWR) (Fig. 1A). Transcriptomic profiling matrices were obtained following the Visium CytAssist Spatial Transcriptomics protocol (Fig. 1B). Subsequent dimensionality reduction and Louvain clustering analysis grouped the spots into distinct clusters (Fig. 1C). These clusters were annotated according to histological structures, encompassing GE with varying GSs and PINs, as well as non-glandular tissues (Fig. 1D). InferCNV, DPT, and PAGA were then performed to evaluate malignancy, progression stages, and developmental trajectories across distinct tissue types (Fig. 1E and F). To identify genes associated with PCa progression, DEG analysis was conducted based on comparisons aligned with GS progression patterns (Fig. 1G). Key genes implicated in PCa development were identified and validated using the TCGA-PRAD dataset (Fig. 1H). Functional interpretation was carried out through GO and KEGG enrichment analyses of these candidate genes. Finally, the differential protein expression of selected genes was confirmed via IHC analysis across representative PCa tissues (Fig. 1I).

The transcriptomic classification of all spatially resolved spots across 12 PCa tissue sections reveals the molecular and spatial heterogeneity of the tissue microenvironment

The transcriptomic expression data from 12 PCa sections were dimensionality reduced using PCA and clustered into distinct groups via Louvain clustering analysis (Fig. 2A; Figs. S1A and S2A). Based on the histological structures observed in HE-stained images, each cluster was pathologically diagnosed and annotated as GSs, PIN, or non-glandular tissues—including fibromuscular stroma, blood vessels

(BVs), and immune infiltration regions (IIRs) (Fig. 2B; Figs. S1B and S2B).

Annotation accuracy was validated by assessing the expression of canonical marker genes specific to each cell type: epithelial cells (TACSTD2, EPCAM), basal cells (KRT5, TP63), luminal cells (KRT8, AR), and PCa cells (KLK3, AMACR, and PCA3)³³ (Fig. 2C; Figs. S1C and S2C). Notably, PCA3 was not detectable in the Visium Cytassist ST data due to limited gene probe coverage and was therefore validated only in the four Visium ST data (Fig. S2C). In the majority of PCa regions across varying GSs, epithelial, luminal, and PCa-specific markers were relatively highly expressed, whereas basal cell markers showed comparatively low expression (Fig. 2C; Figs. S1C and S2C). Conversely, PIN tissues from Cytassist_1, Cytassist_4, and Visium_3 exhibited elevated expression of basal cell markers and reduced levels of PCa-specific markers (Fig. 2C; Fig. S2C), confirming the general concordance between the histological morphology and transcriptomic profiles.

However, certain regions—such as the PIN tissue in Cytassist_2—deviated from expected expression patterns, potentially due to variance-based normalization artifacts or transcriptional contamination from PCa cells mixed within PIN. These discrepancies indicate that marker expression alone may not fully capture histological identity, underscoring the limitations of prior scRNA-seq studies that overrelied on computational annotation without histological correlation.^{21–25}

The malignancy of the GE regions assessed by inferCNV analysis consistent with the GS characteristics

InferCNV analysis has been demonstrated to be an effective tool for predicting clonal hierarchies and assessing tumor malignancy.^{12–16,22,23} The inferCNV values are visualized in the histological images, UMAP embeddings, and chromosomal landscape (Fig. 3A and B; Fig. S3A and S3B). The mean CNV value was used as the primary criterion for assessing the malignancy of the histological clusters.^{22,23} In this study, we computed both the median and mean inferCNV scores for each histological cluster and selected the most appropriate measure based on the distribution characteristics of the data, including normality, skewness, and the presence of outliers. Assessments of tumor malignancy based on the inferCNV index were consistent with histopathological evaluation and showed a general increase with higher GSs across most GE clusters, although not universally (Fig. 3C; Fig. S3C).

Prediction of progression stages and developmental trajectories of GE clusters

We employed DPT and PAGA analyses to infer developmental trajectories and progression stages across distinct clusters.^{56,58} DPT reconstructs cellular developmental progression, identifies branching decisions and differentiation endpoints, and thereby reveals dynamic bifurcation events along differentiation pathways.⁵⁶ Geodesic distances from diffusion maps, combined with dpt_pseudotime values, effectively illustrate both the relational structures and

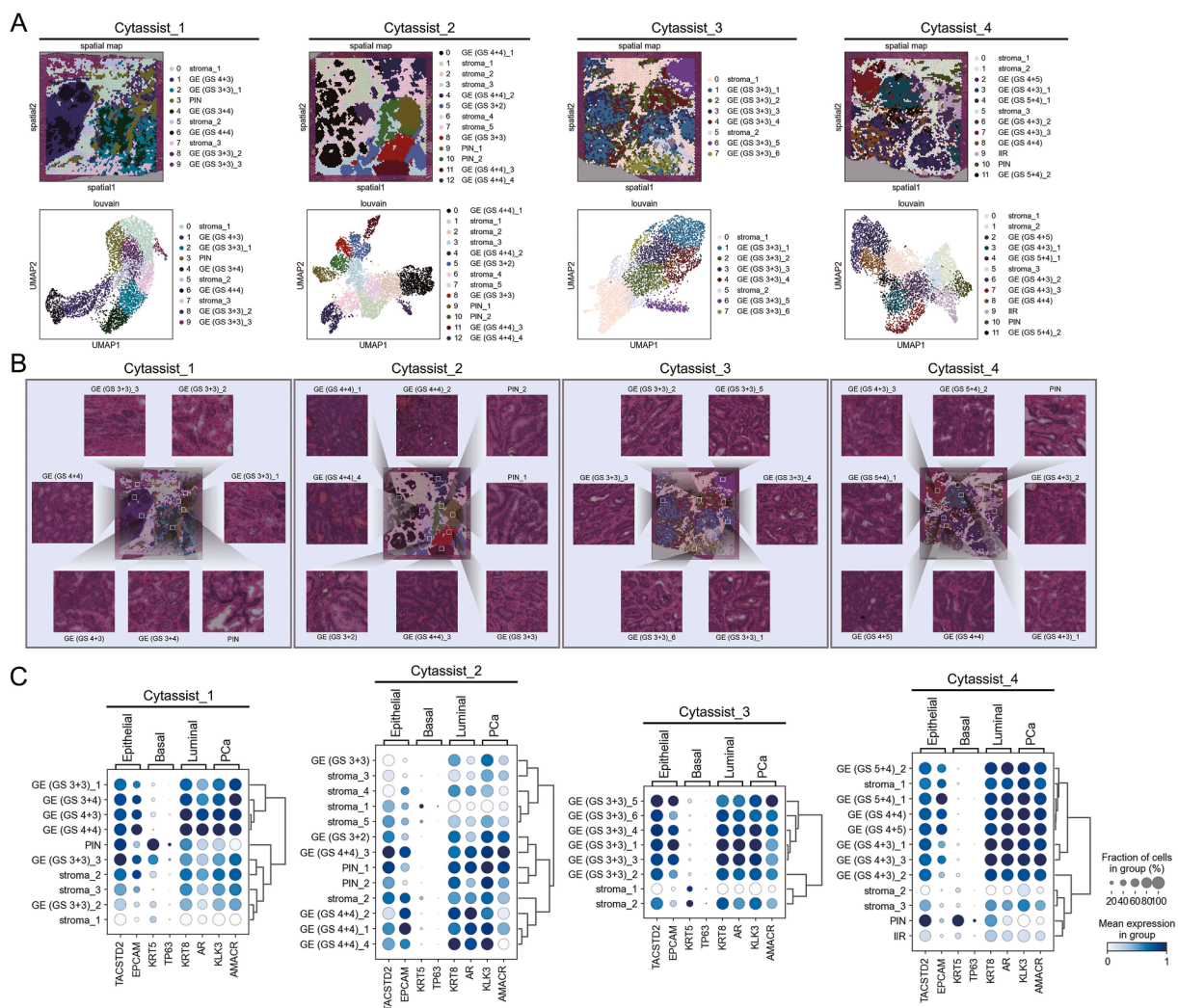


Figure 2 The transcriptomic classification of all spatially resolved spots across CytAssist_1 to CytAssist_4 prostate cancer (PCa) tissue sections into distinct histological structures. **(A)** Transcriptomic classifications across all spots were performed using PCA, UMAP, and Louvain clustering analysis. The resulting clusters were overlaid onto the histological image and annotated according to their corresponding histological structures (upper panel), and visualized in a UMAP embedding (lower panel). **(B)** The regions of distinct clusters were magnified and pathologically evaluated based on corresponding HE-stained images from each PCa section. **(C)** The expression of canonical marker genes—including TACSTD2 and EPCAM (epithelial), KRT5 and TP63 (basal), KRT8 and AR (luminal), and KLK3 and AMACR (PCa)—was analyzed across annotated clusters and visualized as bubble plots.

progression ordering among these clusters (Fig. 4A; Fig. S4A). This trajectory collectively delineates a continuous developmental path and highlights transitional states and lineage relationships between GE clusters. Quantitative assessment indicated a general trend of increasing dpt_pseudotime values with advancing GS, although not universally (Fig. 4B; Fig. S4B).

Complementarily, PAGA captures both discrete disconnected and continuous connected cell-to-cell variations, while also reliably predicting developmental trajectories and gene expression changes from transcriptomic datasets.⁵⁸ The static PAGA graphs, which retain each spot and GE cluster, illustrated the associations and potential developmental trajectories among the GS subtypes and PIN (Fig. 4C; Fig. S4C). To comprehensively evaluate the inter-cluster relationships, degree of malignancy, and progression patterns, we integrated the *cnv_score* and

dpt_pseudotime values into the PAGA graph (Fig. 4D; Fig. S4D). This approach enabled us to identify GE clusters that simultaneously satisfied all three predefined criteria in alignment with GS, thereby establishing a foundation for subsequent DEG analysis according to developmental stage.

DEG analysis based on the developmental progression across varying clusters

DEG analysis was conducted on GE clusters that satisfied all predefined criteria derived from the GS, inferCNV, DPT, and PAGA analyses. Eligible comparisons were required to exhibit consistent relative malignancy levels—as defined by both GS and *cnv_score*—along with coherent progression stages and developmental trajectories within the dpt_pseudotime values and PAGA graphs.

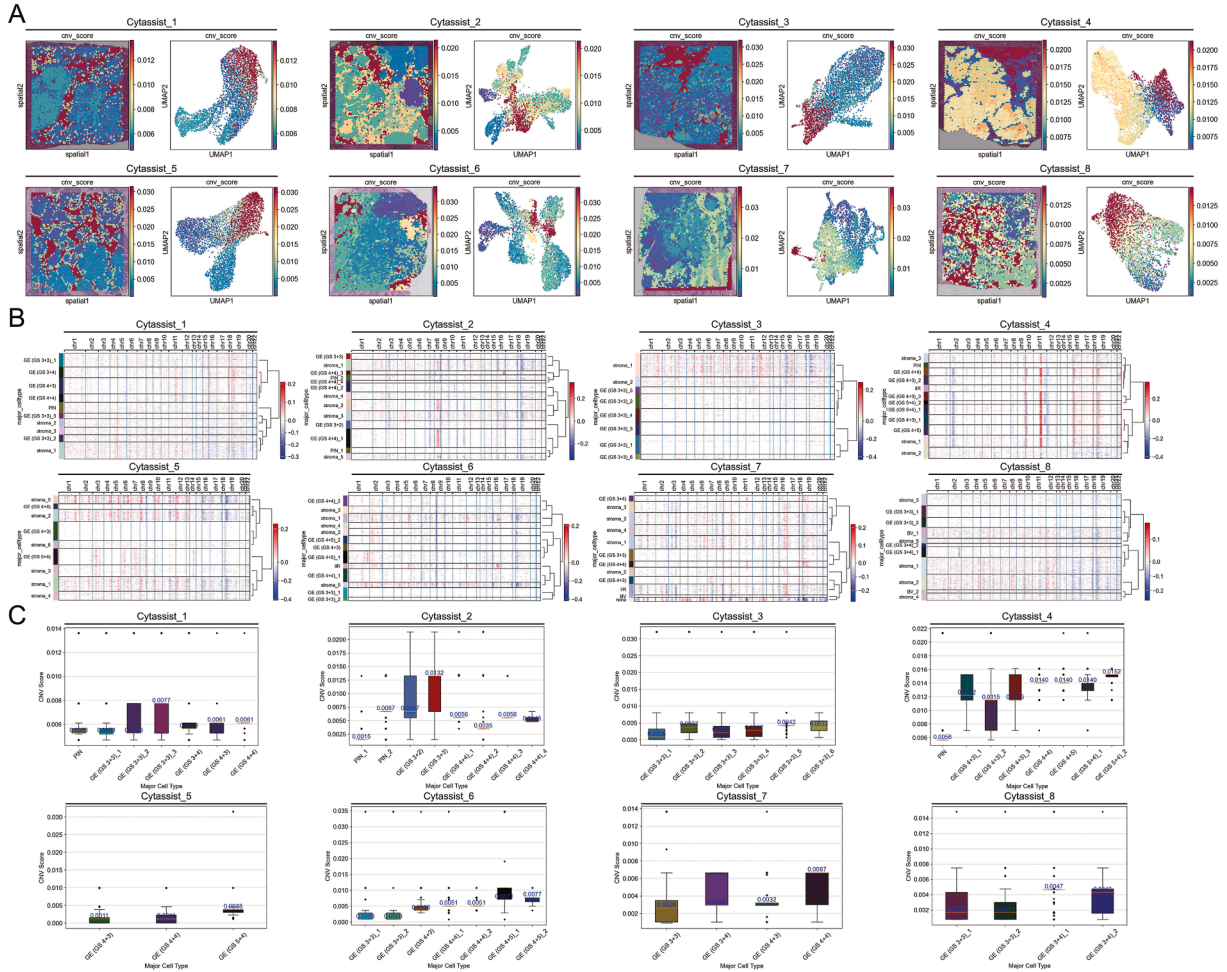


Figure 3 Evaluation of tumor malignancy in varying GE regions across CytAssist_1 to CytAssist_8 prostate cancer (PCa) tissue sections using inferCNV analysis. **(A)** The CNV scores were calculated for spatially resolved spots across eight samples using inferCNV. The scores are visualized as spatial activity maps overlaid on the histological images (left panel) and as UMAP embeddings (right panel). **(B)** The chromosomal landscapes derived from inferCNV values for distinguishing histological clusters across the samples are displayed separately as corresponding heatmaps. **(C)** The median CNV scores, along with their interquartile ranges (IQRs), for GE clusters are summarized using boxplots for each sample.

For example, in CytAssist_1, the PAGA graph revealed strong correlations among GS (3 + 3)_3 and PIN, GS (3 + 4) and GS (3 + 3)_1, and GS (4 + 4) and GS (4 + 3) (Fig. 4C and D). These groups also showed concordant trends in GS, *cnv_score*, and *dpt_pseudotime* values (Figs. 3C and 4B). Based on these consistencies, we included these comparisons in the DEG analysis. Significantly differentially expressed genes were defined as those with $|\ln FC| > 0.25$ and $Q\text{-value} < 0.0001$ (Wilcoxon test) (Fig. 5A). We then compiled genes linked to GS progression that were simultaneously positively (such as *NPY*, *AR*, *FOXA1*, *FOLH1*, and *H2AFJ*) or negatively (such as *JUNB*, *CD74*, *FOS*, and *PTGDS*) regulated across two or three comparisons (Fig. 5B). The same analytical procedure was applied to other samples, with all DEG analyses involving two or three comparative groups (Fig. 3–5; Fig. S3–S5).

In samples such as CytAssist_3 and Visium_4, which exhibited histologically similar GS structures among classified lesions (Fig. 2; Fig. S2), DEG analysis was guided

primarily by developmental progression predictions from inferCNV, DPT, and PAGA analyses (Fig. 5B; Fig. S5B).

Identification of key genes associated with PCa progression across developmental stages

Significant DEGs that appeared in at least two comparisons within each tissue section were identified. These DEGs exhibited consistent patterns of either high or low expression across developmental stages in PCa tissue sections (Fig. 6A; Fig. S7A). The results indicate that among the genes analyzed, *TFF3* was consistently overexpressed in advanced stages across the seven PCa tissues (Fig. 6A). Additionally, several well-established PCa-associated genes, including *OR51E2* (prostate-specific G protein-coupled receptor, PSGR), *FOLH1* (prostate specific membrane antigen, PSMA), *AMACR* (P504S), *FOS* [a subunit of activating protein-1 (AP-1)], *KLK3* (prostate-specific

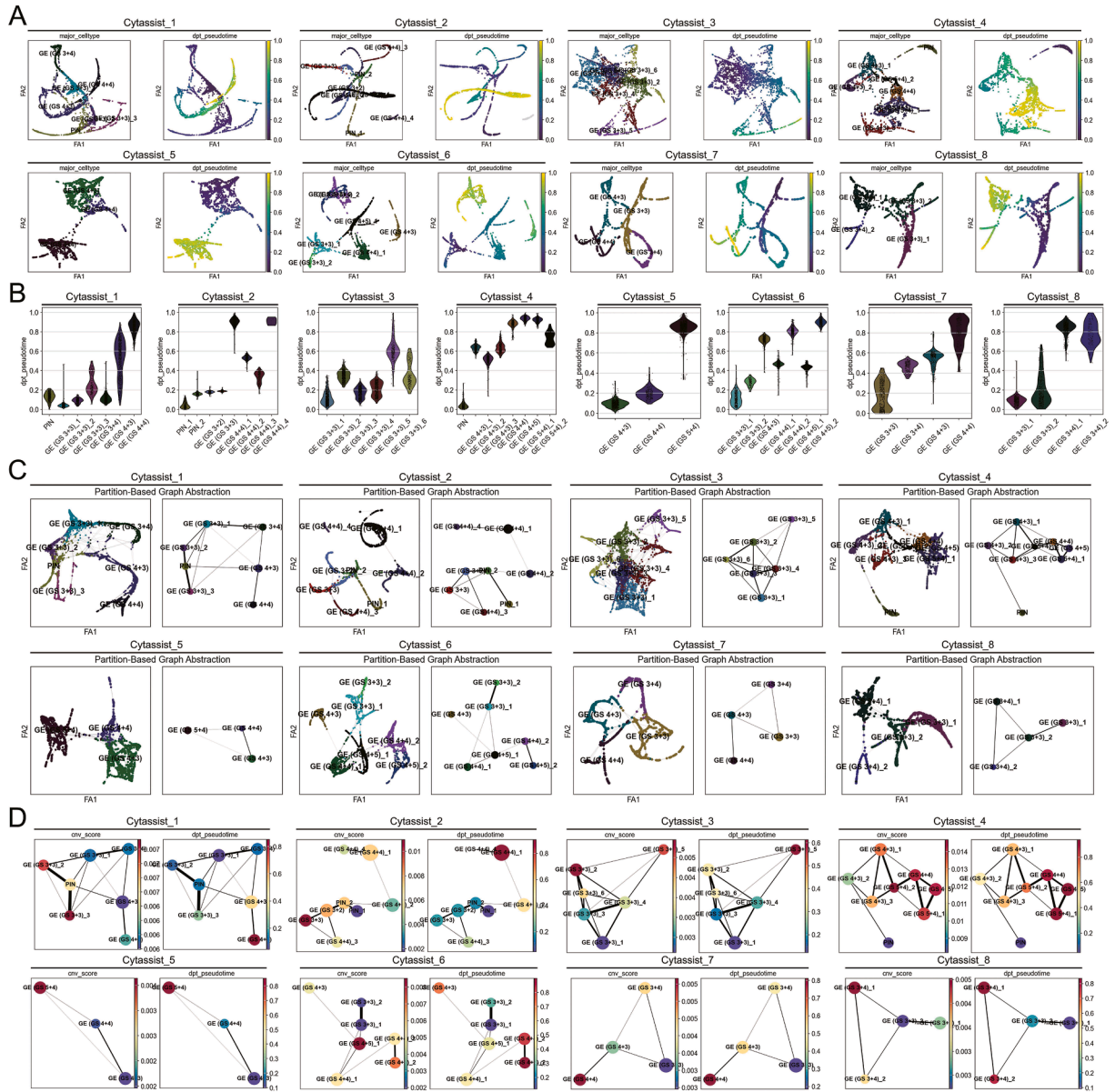


Figure 4 Prediction of progression stages and developmental trajectories across CytAssist_1 to CytAssist_8 prostate cancer (PCa) tissue sections. **(A)** Developmental progression was reconstructed using DPT analysis. Diffusion maps of spatially resolved spots for each sample are presented, displaying coordinates specific to varying GE clusters (left panel) and corresponding dpt_pseudotime values (right panel). **(B)** The quantitative distribution of dpt_pseudotime values within GE clusters for each sample is visualized as violin plots. **(C)** Developmental trajectories of GE clusters for each sample are visualized using PAGA graphs, displaying individual spatial spots (left panel) and GE clusters (right panel). **(D)** The CNV score (left panel) and dpt_pseudotime values (right panel) were integrated into the PAGA graph to illustrate developmental dynamics within each GE cluster.

antigen, PSA), and KLK2, were found to be overexpressed in the advanced stages of at least four PCa tissues (Fig. 6A).

We then identified genes that were consistently overexpressed (9 oncogenes) or downregulated (5 anti-oncogenes) in advanced stages across at least five PCa samples. The expression patterns of these genes were further analyzed within the clusters corresponding to the distinct histological structures of each tissue. In most PCa tissues, these genes exhibited a gradual increase or decrease in expression with GS progression (Fig. 6B and C; Fig. S6A and S6B, Fig. S7B and S7C, Fig. S8A and S8B). Gene

scores derived from the identified 9 oncogenes and 5 anti-oncogenes were calculated and subsequently evaluated across varying GSs within each tissue section. The oncogene scores in most tissues generally increased with GS progression, although not universally (such as CytAssist_5), whereas the scores from anti-oncogenes showed an opposite trend (Fig. 6D; Figs. S6C, S7D, S8C).

Next, we analyzed bilateral lesion samples extracted from the same prostate tissue (CytAssist_3/4 and CytAssist_6/7). Despite originating from the same prostate, each sample pair shared few oncogenes and anti-oncogenes,

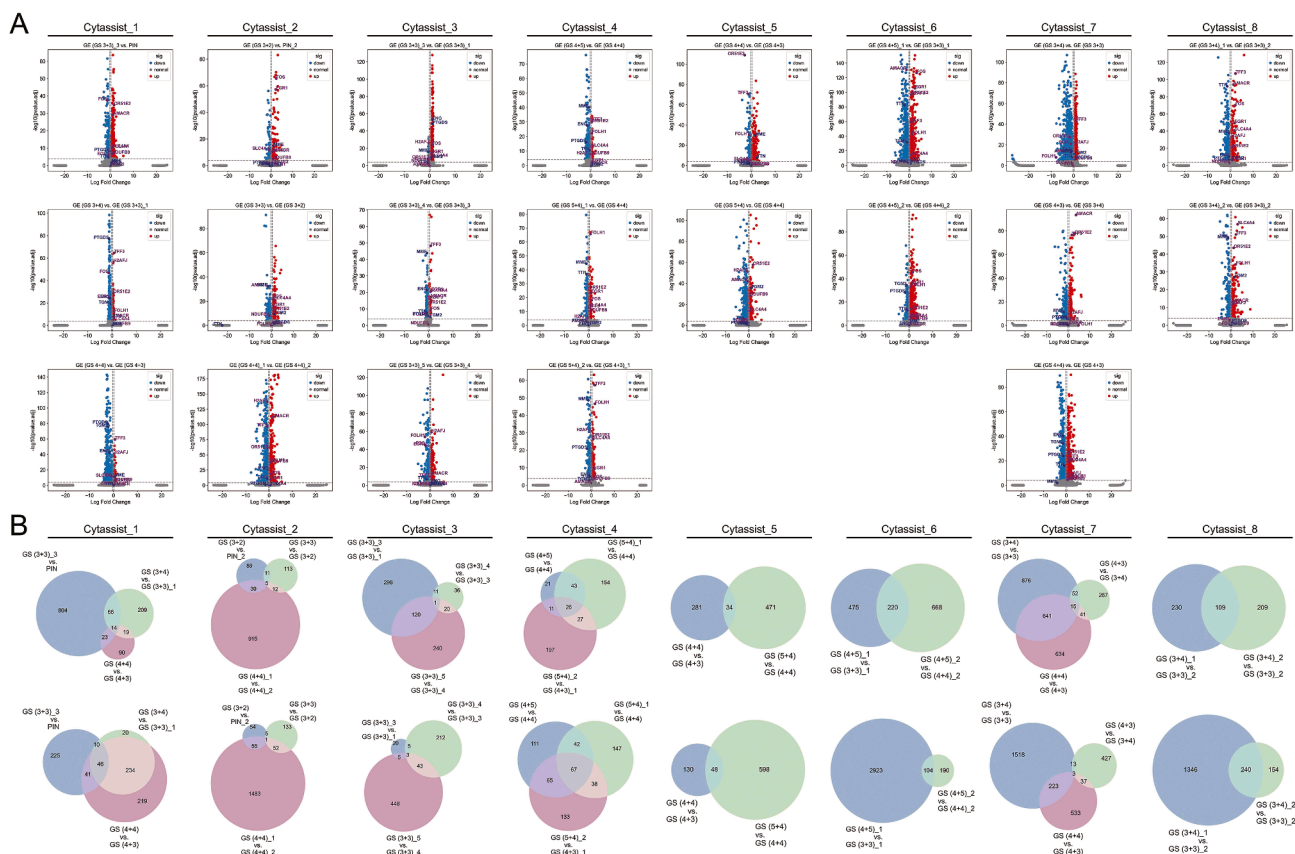


Figure 5 DEG analysis across varying clusters based on progression stages and developmental trajectories in CytAssist_1 to CytAssist_8 prostate cancer (PCa) tissue sections. **(A)** DEG analysis was performed for GE clusters, ensuring that the comparison objects exhibited consistent relative progression stages and developmental trajectories. The comparisons for each tissue section are illustrated using volcano plots, with key genes highlighted: oncogenes (TFF3, OR51E2, FOLH1, AMACR, FOS, SLC4A4, EGR1, NDUFB9, and H2AFJ) in dark magenta and antioncogenes (MME, PTGDS, TTN, ENG, and TGM2) in dark blue. **(B)** The numbers of genes that were significantly and concurrently up-regulated (upper panel) or down-regulated (lower panel) in the above comparisons for each tissue section are depicted in Venn diagrams.

revealing significant transcriptomic heterogeneity in multifocal PCa (Fig. S9A and S9B).

Validation of the identified key genes using the TCGA-PRAD dataset

We subsequently evaluated the genes shared by at least three samples using the TCGA-PRAD dataset. The expression patterns of these genes were analyzed across multiple comparisons based on clinicopathological characteristics, including tumor versus normal tissue, GS > 7 versus GS < 7, pT3 versus pT2, and TP53-mutated versus wild-type tissues. Additionally, survival analyses were conducted to assess the prognostic significance of the gene expression status for RFS and PFS, as previously reported in our studies.^{29,31,32,60,63} Significant DEGs were defined as having a $P < 0.05$ in either the Wilcoxon test or Cox regression analysis.

The results revealed that several oncogenes (such as FOLH1 and NDUFB9) and antioncogenes (such as MME and PTGDS) exhibited generally consistent expression patterns with the ST analysis at advanced stages (Fig. S10A–S10D, S11A–S11D). However, several well-known oncogenes,

including OR51E2 and KLK3, exhibited unexpectedly reduced expression in the majority of advanced states within the TCGA-PRAD dataset (Fig. S10A). Similarly, diverse expression patterns were observed for certain antioncogenes at advanced stages, as exemplified by ENG and TGM2 (Fig. S11A). This discrepancy may arise from the intratumor heterogeneity and histological atypia present in the extracted PCa lesions, which could lead to inconsistencies between the transcriptomic profiles and the pathological diagnoses recorded in the TCGA-PRAD dataset. This comparison underscores that, given the diffusely distributed nature of PCa, ST represents a more appropriate approach for genetic investigation compared to conventional RNA sequencing or related databases.

Furthermore, we evaluated the biological functions of the aforementioned genes. GO and KEGG analyses revealed that oncogenes were significantly enriched in biogenic amine and amine metabolic processes, as well as arginine and proline metabolism, whereas antioncogenes were predominantly associated with extracellular matrix organization and cytoskeleton in muscle cells (Fig. S12A–S12D).

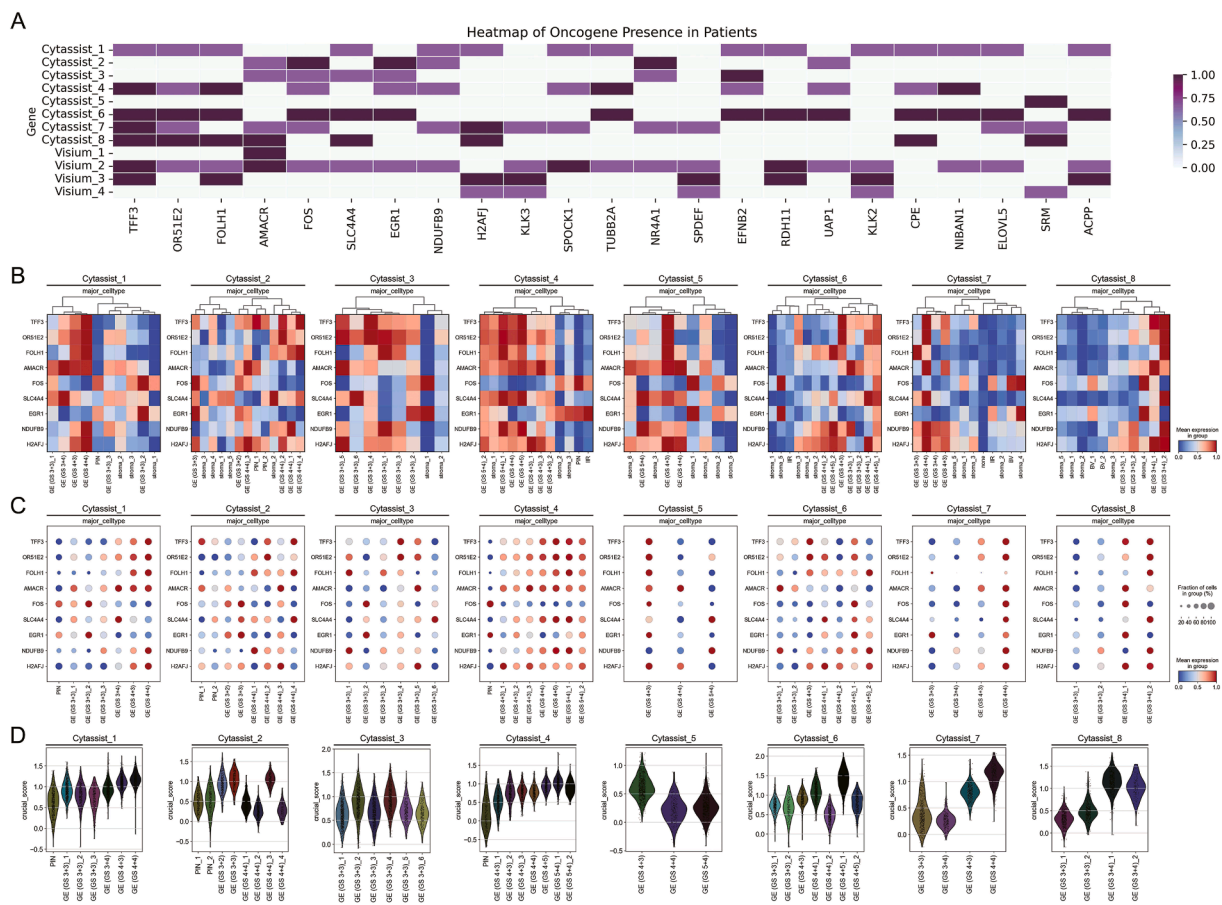


Figure 6 Identifying key oncogenes associated with the progression of developmental stages in 12 prostate cancer (PCa) tissue sections. **(A)** Significant DEGs intersected at least two comparisons within each tissue are summarized and visualized as heatmap. Each gene is assigned a value of 1 (dark purple) if consistently up-regulated across all comparisons within a sample, 0.66 (light purple) if up-regulated in two of three comparisons, and 0 (white) otherwise. **(B)** The expression patterns of oncogenes shared by at least five tissue sections within clusters corresponding to the distinct histological structures of each tissue are visualized as heatmaps. **(C)** The expression patterns of these genes within distinct GE structures of each tissue are displayed as bubble plots. **(D)** Gene scores (crucial_scores) derived from the 9 identified oncogenes were calculated, and their distributions across varying GE clusters in each tissue section are visualized as violin plots.

The proteins encoded by the SLC4A4 and H2AFJ genes exhibit elevated expression levels in advanced-stage PCa tissues

Oncogenes shared by five or more patients were selected as candidates for further validation at the protein level. Among them, several are well-established clinical biomarkers in PCa research, including FOLH1 (PSMA) and AMACR (P504S). Additionally, other shared oncogenes, such as ORS1E2 (PSGR), FOS (a subunit of AP-1), and EGR1 (early growth response protein-1), have also been widely reported in PCa studies. Notably, the association of NDUFB9 with malignancy appeared somewhat tenuous, as it was up-regulated in only two out of the three comparative analyses conducted across all five shared tissue samples.

Based on our screening results and to further investigate novel findings, we subsequently evaluated the protein expression of three candidate genes (TFF3, SLC4A4, and

H2AFJ) by IHC in PIN and PCa tissues obtained from a cohort of 90 patients. The clinicopathological characteristics of these patients are summarized in Table 1. IHC staining revealed a gradual increase in the expression of SLC4A4 and H2AFJ proteins correlating with advancing GSs, whereas TFF3 levels remained unchanged (Fig. 7A). The percentages of positively stained cells and staining indices for these proteins were quantified and compared in PCa tissues stratified by GS categories (GS < 7, GS = 7, and GS > 7) and pT stages (pT2, pT3, and pT4). The results indicated a general increase in the expression of SLC4A4 and H2AFJ proteins with the progression of both the GS and pT stages (Fig. 7B), confirming their potential oncogenic roles in PCa. In contrast, TFF3 protein expression was heterogeneous among GE cells and undetectable in a substantial subset of tissues, which precluded robust statistical analysis (Fig. 7A and B). Nonetheless, its reported overexpression in PCa tissues elsewhere suggests a context-dependent role that warrants further investigation. ^{64–70}

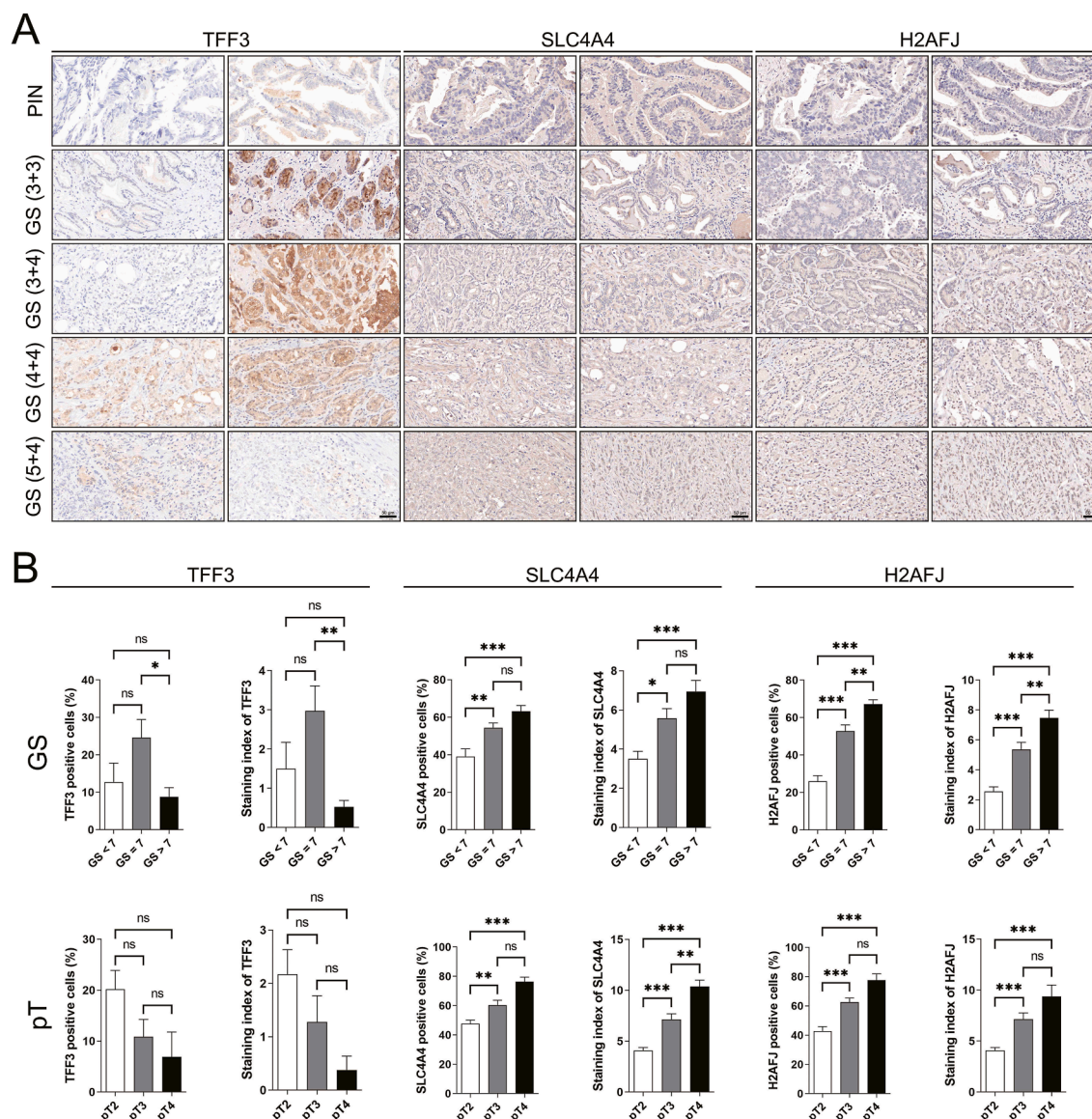


Figure 7 The expression levels of H2AFJ, SLC4A4, and TFF3 proteins in prostate tissues with PIN and varying GSs. **(A)** Representative images of IHC staining for TFF3, SLC4A4, and H2AFJ in prostate cancer (PCa) tissues are presented, illustrating distinct histological classifications, including PIN, GS 3 + 3, GS 3 + 4, GS 4 + 4, and GS 5 + 4. **(B)** The percentages of positively stained cells and the corresponding staining indices for these proteins were quantified and compared across PCa tissues stratified by GSs (GS < 7, GS = 7, and GS > 7) and pT stages (pT2, pT3, and pT4). The results are displayed as histograms, with data represented as the means $\pm$ standard errors of the means (SEMs). ns, not significant. $p > 0.05$; $*p < 0.05$; $**p < 0.01$; and $***p < 0.001$.

Discussion

PCa is a multifocal malignancy originating from GE cells and is characterized by its diffuse distribution throughout prostate tissue,^{2–11} which complicates visual assessment and pathological diagnosis. ST technology is particularly well-suited for PCa research, as it enables high-resolution, spatially resolved quantification of transcriptomic patterns across histologically distinct structures within the prostate.^{16,29,31–36} Previous studies have demonstrated that spatial profiles of factor activity identified through ST closely align with those obtained from IHC staining and histological examination.²⁶ Building on this foundation, the present study aims to precisely identify genes associated

with PCa progression by using ST to analyze DEGs during GS progression in 12 PCa tissue samples.

Analyzing the relationship and origin of multiple lesions within a single sample provides valuable insights into the evolutionary processes underlying PCa. Although current genetic and bioinformatic evidence supports such analyses, it remains insufficient to definitively confirm the multifocal origin of PCa. Some studies have proposed that PCa lesions may originate from independent multiclonal diseases.^{17–20} This hypothesis is supported by analyses of single nucleotide variants (SNVs)¹⁷ and allelic imbalances.^{18–20} In this study, we also identified transcriptomic heterogeneity across bilateral lesions within the same prostate tissue, further underscoring the complexity of PCa development. However,

these prior analyses predominantly focused on a limited subset of prevalent somatic aberrations to infer the dominant clone, which may inadvertently overlook subclonal mutations that are shared among cells due to methodological constraints. Variable patterns of acquired genomic aberrations drive the evolutionary processes of PCa cells, facilitating progression to advanced stages and the emergence of distinct phenotypes within subclones.^{12–16,49,50} High-resolution genome-wide analyses in other studies have demonstrated that PCa cells from distinct foci within a single case trace back to a single genomically aberrant precursor, sharing identical allele-specific CNV changes that are stably propagated through cell division.^{12,13,16} Therefore, these findings support the theory that disparate lesions in multifocal PCa originate from a monoclonal source derived from a single precursor cell,^{12–16} progressing through an ongoing Darwinian evolutionary process.^{71–73}

Heterogeneity in Gleason grade has been widely reported in multifocal PCa.⁷⁴ According to the aforementioned monoclonal theory, the heterogeneity of GSs observed across different regions within a single sample may reflect diverse developmental stages of PCa cells originating from a common monoclonal origin. Transcriptomic separation of multiple lesions through PCA and clustering has been shown to be correlated with various pathological subtypes.^{26,29,31,32} In this study, we applied HVGs for dimensionality reduction via PCA to capture major sources of biological variation. Subsequent UMAP and Louvain clustering were then employed to accurately classify tissue spots into distinct histological structures, as demonstrated in our previous methodology.^{31,32} The resulting classifications were further validated through pathological examination and were annotated based on corresponding histological features.

Although the pathological diagnosis was determined by a senior pathologist, the potential influence of subjective factors necessitates careful consideration to ensure evaluative rigor and reliability. To address this, we computationally validated the accuracy of the annotated histological structures using bioinformatic methods based on both malignancy- and development-related criteria. The mean CNV value (cnv_score) is recognized as a metric for assessing the malignancy of different clusters.^{22,23} Similarly, DPT (dpt_pseudotime) and PAGA are advanced algorithms for pseudotime and trajectory inference, both of which are implemented within the Scanpy framework.^{56–58} In this study, we applied inferCNV, DPT, and PAGA analyses to reconstruct developmental trajectories and progression stages among classified GE regions, including various GSs and PIN. The quantitative results—including cnv_score and dpt_pseudotime, and geodesic distances in the PAGA graph—generally increase with advancing GS, although not uniformly across all samples.

Subsequently, genes associated with PCa progression were identified by performing pairwise comparisons among GE clusters, guided by the preliminary characterization of developmental trajectories and progression stages. The comparison used in this DEG analysis satisfied all criteria for evaluating relative factors across the GS, inferCNV, DPT, and PAGA analyses.

We screened DEGs shared across at least three samples and subsequently evaluated their biological functions

through enrichment analysis. Notably, oncogenes were significantly enriched in pathways related to biogenic amine and amine metabolic processes, as well as arginine and proline metabolism. The identified genes included well-established clinical biomarkers of PCa—such as FOLH1, AMACR, and KLK3—all of which showed consistent association with GS progression in at least five PCa tissue samples.

FOLH1 (prostate-specific membrane antigen, PSMA) is a type II transmembrane glycoprotein overexpressed in PCa, with its expression positively correlated with malignancy, metastatic potential, and early recurrence risk.^{75–77} It serves as a target for PSMA-PET/CT imaging in detecting primary and metastatic lesions.^{78–80} AMACR (P504S) is an IHC marker overexpressed in approximately 90% of PCa cases and high-grade PIN.⁸¹ Its IHC staining is commonly utilized in conjunction with basal cell markers (such as 34βE12 and P63) for diagnostic confirmation.^{82,83} KLK3 (prostate-specific antigen, PSA) is the primary serum biomarker for PCa screening and monitoring, aiding in the early detection and assessment of biochemical recurrence (BCR).^{84–88}

We identified oncogenes, including TFF3, OR51E2, FOLH1, AMACR, FOS, SLC4A4, EGR1, NDUFB9, and H2AFJ, that were consistently expressed in at least five PCa tissue samples. While not yet established as clinical biomarkers, OR51E2, FOS, and EGR1 have been extensively reported in the context of PCa research and analysis.

OR51E2 (prostate-specific G protein-coupled receptor, PSGR) is exclusively expressed in the human prostate and is overexpressed in PCa, with its expression levels correlated with tumor progression.^{89–92} It shows potential as a diagnostic biomarker alongside AMACR and PCA3.^{93,94} Functionally, it promotes inflammation via the interleukin-6 (IL-6) and AKT/NF-κB pathways,^{95–97} though β-ionone-induced activation may suppress proliferation via activation of the MAPK family and inhibition of androgen receptor (AR) nuclear translocation.^{98,99} AP-1 complexes (c-Fos/c-Jun) contribute to PCa progression and therapy resistance.^{100–103} Fos plays dual roles in PCa: it promotes angiogenesis and neuroendocrine (NE) differentiation,¹⁰⁴ yet paradoxically suppresses PCa progression and enhances apoptosis in specific contexts.^{105,106} Although Fos was identified as an oncogene in our ST analysis, its expression was found to be relatively up-regulated in low GS and PIN lesions compared to high GS lesions across Cytassist_1 to Cytassist_5 and Visium_3 samples, underscoring the need for further investigation. EGR1 (early growth response 1) is a key transcription factor that regulates cell growth and survival and promotes PCa progression. It is up-regulated in PCa and associated with high GSs and poor differentiation.^{107,108} It supports androgen-independent growth and metastasis, and its inhibition reduces tumor incidence in models.^{109–112}

The putative oncogene NDUFB9 exhibited inconsistent and biologically tenuous associations with PCa malignancy, as it was only up-regulated in two of the three comparative analyses in all five tissues. Its functional role in PCa remains unexplored. Based on this limited and contextually distinct evidence, we concluded that NDUFB9 is unlikely to play a major role in PCa development.

To identify novel biomarkers for PCa, we performed IHC analysis to validate the expression of proteins encoded by

TFF3, SLC4A4, and H2AFJ in PIN and PCa tissues obtained from 90 patients. The results demonstrated that the protein levels of SLC4A4 and H2AFJ exhibited a progressive increase with GS progression, whereas TFF3 did not exhibit a comparable trend.

TFF3 (trefoil factor 3) is elevated in PCa cells,^{64–70} an effect potentially attributable to promoter hypomethylation.^{69,70} Mechanistically, TFF3 contributes to prostate carcinogenesis by blocking mitochondria-mediated apoptosis.⁶⁴ However, the association between TFF3 expression and GS progression remains controversial. While one study reported elevated TFF3 protein expression in tumors with a GS > 7⁶⁷, others found no correlation or even a negative association with the GS, tumor stage, or recurrence rates.^{65,66,70,113,114} Notably, our IHC validation revealed that TFF3 protein was undetectable in a substantial subset of PCa tissues, including high-grade (GS 5 + 4) cases, complicating robust statistical analysis. Although TFF3 expression has been used for patient stratification to evaluate its progression and prognostic role in PCa, definitive clinical conclusions remain elusive.^{65–67,113,114}

H2AFJ (H2A histone family member J) is significantly overexpressed in luminal epithelial gland cells, including in breast cancer and PCa tissues.¹¹⁵ Similarly, SLC4A4 (solute carrier family 4 member 4) is up-regulated in PCa and promotes disease progression; its inhibition suppresses PCa cell proliferation, migration, invasion, and enzalutamide resistance in preclinical models.^{116–119} Integrated with our findings, these results position SLC4A4 and H2AFJ as promising candidate biomarkers for prostate cancer, warranting further functional and clinical validation.

This study also revealed that well-known oncogenes, including OR51E2, KLK3, and KLK2, which were identified via ST analysis, exhibit negative associations with advanced disease stages in the TCGA-PRAD dataset. This discrepancy underscores inconsistencies between transcriptomic profiles and pathological diagnoses within the TCGA-PRAD, highlighting the potential of ST as a superior predictive tool for genetic research in PCa compared to conventional RNA sequencing and related databases.

Conclusions

In summary, we leveraged ST to comprehensively analyze the transcriptome profiles of 12 PCa samples exhibiting diverse GE histological structures. Through the integration of pairwise comparisons of GE clusters with developmental trajectories and progression stages of classified tumor regions, we identified key genes driving PCa progression. IHC validation further confirmed the elevated expression of SLC4A4 and H2AFJ in advanced GS. Overall, this study establishes an ST-based framework for predicting PCa progression and nominates promising biomarker candidates for clinical translation.

CRedit authorship contribution statement

Yongjun Quan: Writing – original draft, Validation, Supervision, Project administration, Methodology, Investigation, Formal analysis, Data curation. **Mingdong Wang:** Validation, Supervision, Formal analysis, Data curation. **Fan Zou:**

Data curation. **Hong Zhang:** Methodology. **Yishan Zhang:** Validation. **Yongchen Jin:** Supervision. **Hao Ping:** Writing – original draft, Visualization, Validation, Supervision, Funding acquisition, Formal analysis.

Ethics declaration

Ethical approval was granted by the Ethics Committee of Beijing Tongren Hospital, affiliated with Capital Medical University (Approval No.: TREC2023-KY030). Written informed consent was obtained from all participants included in this study.

Data availability

The data and materials supporting the conclusions of this study are included within the article and its supplementary materials. All spatial transcriptomic sequencing data generated in this work have been deposited in the Science Data Bank (ScienceDB). Specifically, the Visium CytAssist ST data can be accessed at <https://cstr.cn/31253.11.sciencedb.25452> with the corresponding DOI: <https://doi.org/10.57760/sciencedb.25452>. The standard Visium ST data are available at <https://cstr.cn/31253.11.sciencedb.09389> and <https://doi.org/10.57760/sciencedb.09389>.

Conflict of interests

The authors declare no competing interests.

Funding

This work was supported by the National Natural Science Foundation of China (No. 82272864 to Hao Ping) and the Capital's Funds for Health Improvement and Research (No. 2024-2-2059 to Hao Ping).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jgendis.2025.101983>.

References

1. Siegel RL, Kratzer TB, Giaquinto AN, Sung H, Jemal A. Cancer statistics. *CA Cancer J Clin.* 2025;75(1):10–45, 2025.
2. Andreou M, Cheng L. Multifocal prostate cancer: biologic, prognostic, and therapeutic implications. *Hum Pathol.* 2010;41(6):781–793.
3. Haffner MC, Zwart W, Roudier MP, et al. Genomic and phenotypic heterogeneity in prostate cancer. *Nat Rev Urol.* 2021;18(2):79–92.
4. Goldstein AS, Huang J, Guo C, Garraway IP, Witte ON. Identification of a cell of origin for human prostate cancer. *Science.* 2010;329(5991):568–571.
5. Wang X, Kruithof-de Julio M, Economides KD, et al. A luminal epithelial stem cell that is a cell of origin for prostate cancer. *Nature.* 2009;461(7263):495–500.
6. Wang ZA, Mitrofanova A, Bergren SK, et al. Lineage analysis of basal epithelial cells reveals their unexpected plasticity and

- supports a cell-of-origin model for prostate cancer heterogeneity. *Nat Cell Biol.* 2013;15(3):274–283.
7. Lawson DA, Zong Y, Memarzadeh S, Xin L, Huang J, Witte ON. Basal epithelial stem cells are efficient targets for prostate cancer initiation. *Proc Natl Acad Sci U S A.* 2010;107(6):2610–2615.
 8. Lee SH, Shen MM. Cell types of origin for prostate cancer. *Curr Opin Cell Biol.* 2015;37:35–41.
 9. Xin L. Cells of origin for cancer: an updated view from prostate cancer. *Oncogene.* 2013;32(32):3655–3663.
 10. Taylor RA, Toivanen R, Frydenberg M, et al. Human epithelial basal cells are cells of origin of prostate cancer, independent of CD133 status. *Stem Cell.* 2012;30(6):1087–1096.
 11. Uchida K, Masumori N, Takahashi A, Itoh N, Tsukamoto T. Characterization of prostatic neuroendocrine cell line established from neuroendocrine carcinoma of transgenic mouse allograft model. *Prostate.* 2005;62(1):40–48.
 12. Boyd LK, Mao X, Xue L, et al. High-resolution genome-wide copy-number analysis suggests a monoclonal origin of multifocal prostate cancer. *Genes Chromosomes Cancer.* 2012;51(6):579–589.
 13. Gundem G, Van Loo P, Kremeyer B, et al. The evolutionary history of lethal metastatic prostate cancer. *Nature.* 2015;520(7547):353–357.
 14. Liu W, Laitinen S, Khan S, et al. Copy number analysis indicates monoclonal origin of lethal metastatic prostate cancer. *Nat Med.* 2009;15(5):559–565.
 15. Hong MK, MacIntyre G, Wedge DC, et al. Tracking the origins and drivers of subclonal metastatic expansion in prostate cancer. *Nat Commun.* 2015;6:6605.
 16. Erickson A, He M, Berglund E, et al. Spatially resolved clonal copy number alterations in benign and malignant tissue. *Nature.* 2022;608(7922):360–367.
 17. Lindberg J, Klevebring D, Liu W, et al. Exome sequencing of prostate cancer supports the hypothesis of independent tumour origins. *Eur Urol.* 2013;63(2):347–353.
 18. Cheng L, Song SY, Pretlow TG, et al. Evidence of independent origin of multiple tumors from patients with prostate cancer. *J Natl Cancer Inst.* 1998;90(3):233–237.
 19. Bostwick DG, Shan A, Qian J, et al. Independent origin of multiple foci of prostatic intraepithelial neoplasia: comparison with matched foci of prostate carcinoma. *Cancer.* 1998;83(9):1995–2002.
 20. Kobayashi M, Ishida H, Shindo T, et al. Molecular analysis of multifocal prostate cancer by comparative genomic hybridization. *Prostate.* 2008;68(16):1715–1724.
 21. Chen S, Zhu G, Yang Y, et al. Single-cell analysis reveals transcriptomic remodellings in distinct cell types that contribute to human prostate cancer progression. *Nat Cell Biol.* 2021;23(1):87–98.
 22. Dong B, Miao J, Wang Y, et al. Single-cell analysis supports a luminal-neuroendocrine transdifferentiation in human prostate cancer. *Commun Biol.* 2020;3(1):778.
 23. Ma X, Guo J, Liu K, et al. Identification of a distinct luminal subgroup diagnosing and stratifying early stage prostate cancer by tissue-based single-cell RNA sequencing. *Mol Cancer.* 2020;19(1):147.
 24. Satija R, Farrell JA, Gennert D, Schier AF, Regev A. Spatial reconstruction of single-cell gene expression data. *Nat Biotechnol.* 2015;33(5):495–502.
 25. Achim K, Pettit JB, Saraiva LR, et al. High-throughput spatial mapping of single-cell RNA-seq data to tissue of origin. *Nat Biotechnol.* 2015;33(5):503–509.
 26. Berglund E, Maaskola J, Schultz N, et al. Spatial maps of prostate cancer transcriptomes reveal an unexplored landscape of heterogeneity. *Nat Commun.* 2018;9(1):2419.
 27. Watanabe R, Miura N, Kurata M, Kitazawa R, Kikugawa T, Saika T. Spatial gene expression analysis reveals characteristic gene expression patterns of *de novo* neuroendocrine prostate cancer coexisting with androgen receptor pathway prostate cancer. *Int J Mol Sci.* 2023;24(10):8955.
 28. Ståhl PL, Salmén F, Vickovic S, et al. Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science.* 2016;353(6294):78–82.
 29. Quan Y, Zhang H, Wang M, Ping H. Visium spatial transcriptomics reveals intratumor heterogeneity and profiles of gleason score progression in prostate cancer. *iScience.* 2023;26(12):108429.
 30. Watanabe R, Miura N, Kurata M, Kitazawa R, Kikugawa T, Saika T. Unveiling the genomic landscape of intraductal carcinoma of the prostate using spatial gene expression analysis. *Int J Mol Sci.* 2024;25(9):4818.
 31. Quan Y, Zhang H, Wang M, Ping H. UQCRB and LBH are correlated with gleason score progression in prostate cancer: spatial transcriptomics and experimental validation. *Comput Struct Biotechnol J.* 2024;23:3315–3326.
 32. Quan Y, Wang M, Zhang H, Lu D, Ping H. Spatial transcriptomics identifies *RBM39* as a gene associated with gleason score progression in prostate cancer. *iScience.* 2024;27(12):111351.
 33. Zou C, Li W, Zhang Y, et al. Identification of an anaplastic subtype of prostate cancer amenable to therapies targeting SP1 or translation elongation. *Sci Adv.* 2024;10(14):eadm7098.
 34. Kiviahio A, Eerola SK, Kallio HML, et al. Single cell and spatial transcriptomics highlight the interaction of club-like cells with immunosuppressive myeloid cells in prostate cancer. *Nat Commun.* 2024;15(1):9949.
 35. Hu J, Liu F, Zhang J, et al. Spatially resolved transcriptomic analysis of the adult human prostate. *Nat Genet.* 2025;57(4):922–933.
 36. Smelik M, Diaz-Roncero Gonzalez D, An X, et al. Combining spatial transcriptomics, pseudotime, and machine learning enables discovery of biomarkers for prostate cancer. *Cancer Res.* 2025;85(13):2514–2526.
 37. Vis AN, Roemeling S, Kranse R, Schröder FH, van der Kwast TH. Should we replace the gleason score with the amount of high-grade prostate cancer? *Eur Urol.* 2007;51(4):931–939.
 38. Gleason DF. Histologic grading of prostate cancer: a perspective. *Hum Pathol.* 1992;23(3):273–279.
 39. Gleason DF. Classification of prostatic carcinomas. *Cancer Chemother Rep.* 1966;50(3):125–128.
 40. Epstein JI, Egevad L, Amin MB, et al. The 2014 international society of urological pathology (ISUP) consensus conference on gleason grading of prostatic carcinoma: definition of grading patterns and proposal for a new grading system. *Am J Surg Pathol.* 2016;40(2):244–252.
 41. Gleason DF, Mellinger GT, Arduino LJ, et al. Prediction of prognosis for prostatic adenocarcinoma by combined histological grading and clinical staging. *J Urol.* 1974;111(1):58–64.
 42. Andrén O, Fall K, Franzén L, Andersson SO, Johansson JE, Rubin MA. How well does the gleason score predict prostate cancer death? A 20-year followup of a population based cohort in Sweden. *J Urol.* 2006;175(4):1337–1340.
 43. Humphrey PA. Gleason grading and prognostic factors in carcinoma of the prostate. *Mod Pathol.* 2004;17(3):292–306.
 44. Lotan TL, Epstein JI. Clinical implications of changing definitions within the gleason grading system. *Nat Rev Urol.* 2010;7(3):136–142.
 45. Popiolek M, Rider JR, Andrén O, et al. Natural history of early, localized prostate cancer: a final report from three decades of follow-up. *Eur Urol.* 2013;63(3):428–435.
 46. Egevad L, Granfors T, Karlberg L, Bergh A, Stattin P. Prognostic value of the gleason score in prostate cancer. *BJU Int.* 2002;89(6):538–542.
 47. Eggener SE, Scardino PT, Walsh PC, et al. Predicting 15-year prostate cancer specific mortality after radical prostatectomy. *J Urol.* 2011;185(3):869–875.

48. Kozminski MA, Tomlins S, Cole A, et al. Standardizing the definition of adverse pathology for lower risk men undergoing radical prostatectomy. *Urol Oncol*. 2016;34(9):415.e1–415.e4156.
49. Ruiz C, Lenkiewicz E, Evers L, et al. Advancing a clinically relevant perspective of the clonal nature of cancer. *Proc Natl Acad Sci U S A*. 2011;108(29):12054–12059.
50. Robinson D, Van Allen EM, Wu YM, et al. Integrative clinical genomics of advanced prostate cancer. *Cell*. 2015;161(5):1215–1228.
51. Ross HM, Kryvenko ON, Cowan JE, Simko JP, Wheeler TM, Epstein JI. Do adenocarcinomas of the prostate with gleason score (GS) ≤ 6 have the potential to metastasize to lymph nodes? *Am J Surg Pathol*. 2012;36(9):1346–1352.
52. Donin NM, Laze J, Zhou M, Ren Q, Lepor H. Gleason 6 prostate tumors diagnosed in the PSA era do not demonstrate the capacity for metastatic spread at the time of radical prostatectomy. *Urology*. 2013;82(1):148–152.
53. Kweldam CF, Wildhagen MF, Bangma CH, van Leenders GJ. Disease-specific death and metastasis do not occur in patients with gleason score ≤ 6 at radical prostatectomy. *BJU Int*. 2015;116(2):230–235.
54. Eggener SE, Berlin A, Vickers AJ, Paner GP, Wolinsky H, Cooperberg MR. Low-grade prostate cancer: time to stop calling it cancer. *J Clin Oncol*. 2022;40(27):3110–3114.
55. Labbate CV, Paner GP, Eggener SE. Should grade group 1 (GG1) be called cancer? *World J Urol*. 2022;40(1):15–19.
56. Haghverdi L, Büttner M, Wolf FA, Büttner F, Theis FJ. Diffusion pseudotime robustly reconstructs lineage branching. *Nat Methods*. 2016;13(10):845–848.
57. Wolf FA, Angerer P, Theis FJ. SCANPY large-scale single-cell gene expression data analysis. *Genome Biol*. 2018;19(1):15.
58. Wolf FA, Hamey FK, Plass M, et al. PAGA: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. *Genome Biol*. 2019;20(1):59.
59. Liu J, Lichtenberg T, Hoadley KA, et al. An integrated TCGA pan-cancer clinical data resource to drive high-quality survival outcome analytics. *Cell*. 2018;173(2):400–416.e11.
60. Quan Y, Zhang X, Ping H. Construction of a risk prediction model using m6A RNA methylation regulators in prostate cancer: comprehensive bioinformatic analysis and histological validation. *Cancer Cell Int*. 2022;22(1):33.
61. Wang L, Li J, Hou J, et al. p53 expression but not p16(INK4A) correlates with human papillomavirus-associated esophageal squamous cell carcinoma in Kazakh population. *Infect Agent Cancer*. 2016;11:19.
62. Wang L, Zhang Z, Yu X, et al. Unbalanced YAP-SOX9 circuit drives stemness and malignant progression in esophageal squamous cell carcinoma. *Oncogene*. 2019;38(12):2042–2055.
63. Quan Y, Zhang X, Wang M, Ping H. Histone lysine methylation patterns in prostate cancer microenvironment infiltration: integrated bioinformatic analysis and histological validation. *Front Oncol*. 2022;12:981226.
64. Liu J, Kim SY, Shin S, et al. Overexpression of TFF3 is involved in prostate carcinogenesis via blocking mitochondria-mediated apoptosis. *Exp Mol Med*. 2018;50(8):1–11.
65. Garraway IP, Seligson D, Said J, Horvath S, Reiter RE. Trefoil factor 3 is overexpressed in human prostate cancer. *Prostate*. 2004;61(3):209–214.
66. Faith DA, Isaacs WB, Morgan JD, et al. Trefoil factor 3 overexpression in prostatic carcinoma: prognostic importance using tissue microarrays. *Prostate*. 2004;61(3):215–227.
67. Timofte AD, Giuscă SE, Lozneanu L, et al. HOXB13 and TFF3 can contribute to the prognostic stratification of prostate adenocarcinoma. *Rom J Morphol Embryol*. 2021;62(1):41–52.
68. Adler D, Lindström A, Büttner R, Wernert N. Analysis of laser-microdissected prostate cancer tissues reveals potential tumor markers. *Int J Mol Med*. 2011;28(4):605–611.
69. Vestergaard EM, Nexø E, Tørring N, Borre M, Ørntoft TF, Sørensen KD. Promoter hypomethylation and upregulation of trefoil factors in prostate cancer. *Int J Cancer*. 2010;127(8):1857–1865.
70. Nørgaard M, Haldrup C, Storebjerg TM, et al. Comprehensive evaluation of TFF3 promoter hypomethylation and molecular biomarker potential for prostate cancer diagnosis and prognosis. *Int J Mol Sci*. 2017;18(9):2017.
71. Nowell PC. The clonal evolution of tumor cell populations. *Science*. 1976;194(4260):23–28.
72. Greaves M, Maley CC. Clonal evolution in cancer. *Nature*. 2012;481(7381):306–313.
73. Gerlinger M, Rowan AJ, Horswell S, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N Engl J Med*. 2012;366(10):883–892.
74. Arora R, Koch MO, Eble JN, Ulbright TM, Li L, Cheng L. Heterogeneity of gleason grade in multifocal adenocarcinoma of the prostate. *Cancer*. 2004;100(11):2362–2366.
75. Silver DA, Pellicer I, Fair WR, Heston WD, Cordon-Cardo C. Prostate-specific membrane antigen expression in normal and malignant human tissues. *Clin Cancer Res*. 1997;3(1):81–85.
76. Ross JS, Sheehan CE, Fisher HA, et al. Correlation of primary tumor prostate-specific membrane antigen expression with disease recurrence in prostate cancer. *Clin Cancer Res*. 2003;9(17):6357–6362.
77. Sweat SD, Pacelli A, Murphy GP, Bostwick DG. Prostate-specific membrane antigen expression is greatest in prostate adenocarcinoma and lymph node metastases. *Urology*. 1998;52(4):637–640.
78. Hofman MS, Lawrentschuk N, Francis RJ, et al. Prostate-specific membrane antigen PET-CT in patients with high-risk prostate cancer before curative-intent surgery or radiotherapy (proPSMA): a prospective, randomised, multicentre study. *Lancet*. 2020;395(10231):1208–1216.
79. Perera M, Papa N, Christidis D, et al. Sensitivity, specificity, and predictors of positive ^{68}Ga -prostate-specific membrane antigen positron emission tomography in advanced prostate cancer: a systematic review and meta-analysis. *Eur Urol*. 2016;70(6):926–937.
80. Pienta KJ, Gorin MA, Rowe SP, et al. A phase 2/3 prospective multicenter study of the diagnostic accuracy of prostate specific membrane antigen PET/CT with ^{18}F -DCFPyL in prostate cancer patients (OSPREY). *J Urol*. 2021;206(1):52–61.
81. Rubin MA, Zhou M, Dhanasekaran SM, et al. Alpha-Methylacyl coenzyme A racemase as a tissue biomarker for prostate cancer. *JAMA*. 2002;287(13):1662–1670.
82. Epstein JI, Egevad L, Humphrey PA, Montironi R, Members of the ISUP Immunohistochemistry in Diagnostic Urologic Pathology Group. Best practices recommendations in the application of immunohistochemistry in the prostate: report from the International Society of Urologic Pathology consensus conference. *Am J Surg Pathol*. 2014;38(8):e6–e19.
83. Box A, Alshalalfa M, Hegazy SA, Donnelly B, Bismar TA. High alpha-methylacyl-coa racemase (AMACR) is associated with ERG expression and with adverse clinical outcome in patients with localized prostate cancer. *Tumor Biol*. 2016;37(9):12287–12299.
84. Osses DF, Remmers S, Schröder FH, van der Kwast T, Roobol MJ. Results of prostate cancer screening in a unique cohort at 19yr of follow-up. *Eur Urol*. 2019;75(3):374–377.
85. Hugosson J, Roobol MJ, Månsson M, et al. A 16-yr follow-up of the European randomized study of screening for prostate cancer. *Eur Urol*. 2019;76(1):43–51.
86. Hugosson J, Godtman RA, Carlsson SV, et al. Eighteen-year follow-up of the göteborg randomized population-based prostate cancer screening trial: effect of sociodemographic variables on participation, prostate cancer incidence and mortality. *Scand J Urol*. 2018;52(1):27–37.

87. Grivas N, de Bruin D, Barwari K, et al. Ultrasensitive prostate-specific antigen level as a predictor of biochemical progression after robot-assisted radical prostatectomy: towards risk adapted follow-up. *J Clin Lab Anal.* 2019;33(2):e22693.
88. Roach M, Hanks G, Thames Jr H, et al. Defining biochemical failure following radiotherapy with or without hormonal therapy in men with clinically localized prostate cancer: recommendations of the RTOG-ASTRO Phoenix Consensus conference. *Int J Radiat Oncol Biol Phys.* 2006;65(4):965–974.
89. Weng J, Wang J, Cai Y, et al. Increased expression of prostate-specific G-protein-coupled receptor in human prostate intraepithelial neoplasia and prostate cancers. *Int J Cancer.* 2005;113(5):811–818.
90. Xia C, Ma W, Wang F, Hua Sb, Liu M. Identification of a prostate-specific G-protein coupled receptor in prostate cancer. *Oncogene.* 2001;20(41):5903–5907.
91. Xu LL, Stackhouse BG, Florence K, et al. *PSGR* a novel prostate-specific gene with homology to a G protein-coupled receptor, is overexpressed in prostate cancer. *Cancer Res.* 2000;60(23):6568–6572.
92. Xu LL, Sun C, Petrovics G, et al. Quantitative expression profile of *PSGR* in prostate cancer. *Prostate Cancer Prostatic Dis.* 2006;9(1):56–61.
93. Wang J, Weng J, Cai Y, Penland R, Liu M, Ittmann M. The prostate-specific G-protein coupled receptors *PSGR* and *PSGR2* are prostate cancer biomarkers that are complementary to alpha-methylacyl-coa racemase. *Prostate.* 2006;66(8):847–857.
94. Rigau M, Morote J, Mir MC, et al. *PSGR* and *PCA3* as biomarkers for the detection of prostate cancer in urine. *Prostate.* 2010;70(16):1760–1767.
95. Rodriguez M, Luo W, Weng J, et al. *PSGR* promotes prostatic intraepithelial neoplasia and prostate cancer xenograft growth through NF- κ B. *Oncogenesis.* 2014;3(8):e114.
96. Rodriguez M, Siwko S, Zeng L, Li J, Yi Z, Liu M. Prostate-specific G-protein-coupled receptor collaborates with loss of *PTEN* to promote prostate cancer progression. *Oncogene.* 2016;35(9):1153–1162.
97. Weng J, Ma W, Mitchell D, Zhang J, Liu M. Regulation of human prostate-specific G-protein coupled receptor, *PSGR*, by two distinct promoters and growth factors. *J Cell Biochem.* 2005;96(5):1034–1048.
98. Neuhaus EM, Zhang W, Gelis L, Deng Y, Noldus J, Hatt H. Activation of an olfactory receptor inhibits proliferation of prostate cancer cells. *J Biol Chem.* 2009;284(24):16218–16225.
99. Xie H, Liu T, Chen J, et al. Activation of *PSGR* with β -ionone suppresses prostate cancer progression by blocking androgen receptor nuclear translocation. *Cancer Lett.* 2019;453:193–205.
100. Quan Y, Zhang X, Butler W, et al. The role of N-cadherin/c-Jun/NDRG1 axis in the progression of prostate cancer. *Int J Biol Sci.* 2021;17(13):3288–3304.
101. Edwards J, Krishna NS, Mukherjee R, Bartlett JM. The role of c-Jun and c-Fos expression in androgen-independent prostate cancer. *J Pathol.* 2004;204(2):153–158.
102. Ouyang X, Jessen WJ, Al-Ahmadie H, et al. Activator protein-1 transcription factors are associated with progression and recurrence of prostate cancer. *Cancer Res.* 2008;68(7):2132–2144.
103. Wang Z, Townley SL, Zhang S, et al. *FOXA2* rewires AP-1 for transcriptional reprogramming and lineage plasticity in prostate cancer. *Nat Commun.* 2024;15(1):4914.
104. Collado B, Sánchez MG, Díaz-Laviada I, Prieto JC, Carmena MJ. Vasoactive intestinal peptide (VIP) induces c-fos expression in LNCaP prostate cancer cells through a mechanism that involves Ca^{2+} signalling. Implications in angiogenesis and neuroendocrine differentiation. *Biochim Biophys Acta.* 2005;1744(2):224–233.
105. Riedel M, Berthelsen MF, Cai H, et al. *In vivo* CRISPR inactivation of fos promotes prostate cancer progression by altering the associated AP-1 subunit Jun. *Oncogene.* 2021;40(13):2437–2447.
106. Zhang X, Zhang L, Yang H, et al. C-Fos as a proapoptotic agent in TRAIL-induced apoptosis in prostate cancer cells. *Cancer Res.* 2007;67(19):9425–9434.
107. Thigpen AE, Cala KM, Guileyardo JM, Molberg KH, McConnell JD, Russell DW. Increased expression of early growth response-1 messenger ribonucleic acid in prostatic adenocarcinoma. *J Urol.* 1996;155(3):975–981.
108. Eid MA, Kumar MV, Iczkowski KA, Bostwick DG, Tindall DJ. Expression of early growth response genes in human prostate cancer. *Cancer Res.* 1998;58(11):2461–2468.
109. Baron V, De Gregorio G, Krones-Herzig A, et al. Inhibition of Egr-1 expression reverses transformation of prostate cancer cells *in vitro* and *in vivo*. *Oncogene.* 2003;22(27):4194–4204.
110. Yang SZ, Eltoum IA, Abdulkadir SA. Enhanced EGR1 activity promotes the growth of prostate cancer cells in an androgen-depleted environment. *J Cell Biochem.* 2006;97(6):1292–1299.
111. Ma Y, Cheng Q, Ren Z, et al. Induction of IGF-1R expression by EGR-1 facilitates the growth of prostate cancer cells. *Cancer Lett.* 2012;317(2):150–156.
112. Li L, Ameri AH, Wang S, et al. EGR1 regulates angiogenic and osteoclastogenic factors in prostate cancer and promotes metastasis. *Oncogene.* 2019;38(35):6241–6255.
113. Abou-Ouf H, Ghosh S, Box A, Palanisamy N, Bismar TA. Combined loss of TFF3 and PTEN is associated with lethal outcome and overall survival in men with prostate cancer. *J Cancer Res Clin Oncol.* 2019;145(7):1751–1759.
114. Al Bashir S, Alorjani MS, Kheirallah K, et al. PTEN, ERG, SPINK1, and TFF3 status and relationship in a prostate cancer cohort from Jordanian Arab population. *Medicina (Kaunas).* 2024;60(1):174.
115. Redon CE, Schmal Z, Tewary G, et al. Histone variant H2A.J is enriched in luminal epithelial gland cells. *Genes.* 2021;12(11):1665.
116. Liu Z, Wang Q, Zhai G, Ke S, Yu X, Guo J. SLC4A4 promotes prostate cancer progression *in vivo* and *in vitro* via AKT-mediated signalling pathway. *Cancer Cell Int.* 2022;22(1):127.
117. Xiao F, Liu J, Zheng Y, et al. The targeted inhibition of prostate cancer by iron-based nanoparticles based on bioinformatics. *J Biomater Appl.* 2021;36(1):3–14.
118. Rashid A, Fung HL, Tang AHN. Dissecting the novel molecular interactions of solute carrier family 4 member 4 (SLC4A4) for prostate cancer (PCa) progression. *Sci Rep.* 2024;14(1):29133.
119. Chen HJ, Yu MM, Huang JC, et al. SLC4A4 is a novel driver of enzalutamide resistance in prostate cancer. *Cancer Lett.* 2024;597:217070.