Molecular Genetics and Genomics (2026) 301:50
https://doi.org/10.1007/s00438-026-02371-w

ORIGINAL ARTICLE

Single-cell transcriptomic profiling combined with Mendelian randomization illuminates molecular drivers of bladder cancer

Junrui He[1,6] · Zhuoying Jiang[2,8] · Shan Peng[3,7] · Lijuan Peng[2,8] · Zhongyou Xia[4,6] · Ting Ma[5] · Xiaojun Tan[4,6] · Qiongxian Long[2,8] · Ji Wu[4,6]

Received: 29 July 2025 / Accepted: 26 January 2026

Abstract

Bladder cancer exhibits marked cellular heterogeneity, which complicates the identification of molecular drivers with both functional and genetic relevance. The link between genetic susceptibility and transcriptional activity in tumorigenesis remains incompletely understood. Single-cell transcriptomes of bladder carcinoma and adjacent normal tissues were processed using Seurat and Harmony for data normalization, integration, and dimensionality reduction. Differentially expressed genes were further prioritized using causal inference based on genome-wide association and expression quantitative trait loci data. Functional analyses included pathway enrichment, immune profiling, and transcription factor network inference. Candidate genes were validated through immunohistochemistry, quantitative PCR, and Western blotting, and in vitro cell-based assays (including CCK-8, scratch wound healing, transwell invasion, and flow cytometry analysis). Among all cellular populations, B cells showed the highest fold-change score. Mendelian randomization highlighted six bladder-cancer-related genes—*ARHGEF18*, *HLA-DRB5*, *ISG20*, *NCF1*, *RPL13*, and *YPEL5*. *ARHGEF18* raised risk (OR 1.001; 95% CI 1.000–1.002; P 0.017), whereas *YPEL5* showed a protective association (OR 0.997; 95% CI 0.995–0.999; P 0.003). Immunohistochemistry confirmed elevated *ARHGEF18* and reduced *YPEL5* expression in tumor tissues. In vitro experiments demonstrated that *ARHGEF18* knockdown and *YPEL5* overexpression led to decreased cell proliferation, migration, invasion, and increased apoptosis in T24 and 5637 bladder cancer cells. Our findings identify *ARHGEF18* and *YPEL5* as genetically and transcriptionally supported regulators of bladder cancer, highlighting a scalable strategy for linking genetic risk to cell-type–specific mechanisms.

Keywords Bladder cancer · Mendelian randomization · Single-cell transcriptomics · Immune profiling · Immunometabolism

Communicated by Wenfei Jin

✉ Ji Wu
wuji2168@163.com

[1] Department of Urology, Nanchong Central Hospital Affiliated to North Sichuan Medical College, Nanchong 637000, China

[2] Department of Pathology, Nanchong Central Hospital Affiliated to North Sichuan Medical College, Nanchong 637000, China

[3] Department of Medical Rehabilitation, Beijing Anzhen Nanchong Hospital, Capital Medical University & Nanchong Central Hospital, Nanchong 637000, China

[4] Department of Urology, Beijing Anzhen Nanchong Hospital, Capital Medical University & Nanchong Central Hospital, Nanchong 637000, China

[5] Department of Clinical Medicine, North Sichuan Medical College, Nanchong 637000, China

[6] Department of Urology, The Second Clinical Medical College of North, Sichuan Medical College, Nanchong 637000, China

[7] Department of Medical Rehabilitation, The Second Clinical Medical College of North Sichuan Medical College, Nanchong 637000, China

[8] Department of Pathology, The Second Clinical Medical College of North, Sichuan Medical College, Nanchong 637000, China

Published online: 04 March 2026

Abbreviations

GWAS	Genome-wide association studies
scRNA-seq	Single-cell RNA sequencing
MR	Mendelian randomization
NCBI	National center for biotechnology information
SNP	Single nucleotide polymorphism
MAD	Median absolute deviation
UMAP	Unified manifold approximation and projection
GO	Gene ontology
IVs	Instrumental variables
IVW	Inverse variance weighted
LD	Linkage disequilibrium
InSIDE	Independent of direct effects
GSEA	Gene set enrichment analysis
GSVA	Gene set variation analysis
NES	Normalized enrichment score
AUC	Area under the curve
IHC	Immunohistochemistry
qPCR	Quantitative real-time polymerase chain reaction
WB	Western blot
SD	Standard deviation
IQR	Interquartile range
TAMs	Tumor-associated macrophages

Introduction

Bladder cancer is a biologically complex malignancy in which genetic risk, cellular heterogeneity, and tumor–immune interactions jointly shape disease progression and therapeutic response, and it remains one of the top ten most prevalent cancers worldwide, with marked regional and demographic disparities driven in part by aging and tobacco exposure (Jubber et al. 2023). In the year 2020, global data indicated around 573,000 new diagnoses and approximately 213,000 fatalities attributed to this disease (van Hoogstraten et al. 2023). Despite progress in surgical methods, chemotherapy, and immunotherapy, managing advanced Bladder cancer—especially in cases that are muscle-invasive and metastatic—remains a significant challenge (Kim and Lee 2020; Patel et al. 2020; Stein et al. 2023; Lerner et al. 2024; Boll et al. 2025). Such challenges stem from the high heterogeneity of tumors and limited insight into the molecular processes underlying their initiation and progression.

Genome-wide association studies (GWAS) have identified several bladder cancer-associated loci, including regions at 4p16.3, 5p15.33, 5q12.3, 8q21.13, and 19q13.33. Despite this, the precise causal genes and their functions in tumor development are still not well comprehended, which restricts their use in translational potential (Rafnar et al. 2011; Wang et al. 2016; Koutros et al. 2023). Additionally, although somatic mutations in genes like *TP53*, *FGFR3*, *PIK3CA*, and *RB1*, as well as epigenetic regulators such as *KDM6A*, have been identified as elements in bladder carcinoma biology, such discoveries frequently overlook the influence of non-coding genomic elements and the complex cellular composition within the tumor milieu (Barrows et al. 2020; Essakly et al. 2020; Komura et al. 2023; Zhu et al. 2024). This gap in knowledge underscores the imperative to elucidate the cell-type-specific regulatory mechanisms involved.

Recent advancements in methodology now facilitate a more accurate exploration of these unanswered questions. Single-cell RNA sequencing (scRNA-seq) provides exceptional resolution for delineating tumor heterogeneity, while Mendelian randomization (MR) offers a solid foundation for determining causal links between genetic variants and gene expression (Hu et al. 2024; Kim et al. 2024). Integrating scRNA-seq with MR has been applied in both cancer and non-cancer research, enabling the identification of cell type-specific genes with causal roles and helping link genetic risk factors to disease mechanisms (Chen et al. 2024b; Bai et al. 2025; Gong et al. 2025; Wu et al. 2025).

In this research, we utilize a comprehensive method that merges scRNA-seq with MR to systematically identify genes associated with Bladder cancer in defined cellular environments. This approach enhances our insight into bladder cancer biology and lays a foundation for creating more accurate diagnostic and therapeutic strategies.

Materials and methods

Data acquisition

(1) The Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/), maintained by the National Center for Biotechnology Information (NCBI; https://www.ncbi.nlm.nih.gov), serves as a public repository for gene expression data. 13 samples from GSE222315, including four normal and nine tumor tissues, were obtained from the NCBI GEO repository for single-cell analysis (Zheng et al. 2024). In addition, three GEO datasets were retrieved for bulk transcriptomic validation: GSE121711 (8 tumor, 10 normal) (Loras et al. 2019), GSE13507 (165 tumor, 67 normal) (Lee et al. 2010), and GSE7476 (9 tumor, 3 normal) (Mengual et al. 2009).
(2) The Cancer Genome Atlas (TCGA; https://portal.gdc.cancer.gov/) is the largest publicly available can

cer genomics resource, offering multi-omic datasets—including transcriptomic profiles, copy-number variations, and single-nucleotide polymorphism calls (SNPs). Within this public database, mRNA expression information was collected for 431 patients diagnosed with bladder cancer (TCGA-BLCA), which comprises 19 normal samples and 412 tumor samples.

(3) Exposure data: Blood-derived eQTL data were downloaded from the eQTLGen consortium (https://www.eqtlgen.org), which investigates how genetic variation governs transcript levels in blood and contributes to complex traits; its ongoing Phase II expands this effort through genome-wide meta-analyses.

(4) Outcome data: This study utilized GWAS summary data derived from European-ancestry participants, obtained via the IEU OpenGWAS database (ieu-b-4874; https://opengwas.io). The GWAS Catalog, which aggregates studies, key variant-disease links, and complete summary statistics mapped to reference genome builds and dbSNP versions, contains data from 1,279 bladder cancer cases and 372,016 controls.

Quality control

The expression matrix was imported using Seurat V4. Cells were subsequently filtered based on total UMI counts, gene detection levels, and the relative abundance of mitochondrial and ribosomal transcripts, indicating their contribution to overall gene expression. Elevated levels of mitochondrial and ribosomal gene expression imply poor RNA quality and a risk of cell apoptosis. Outliers were detected by applying a median absolute deviation (MAD) threshold set at three times the median. Following this, potential doublets were removed using DoubletFinder (V2.0.4), completing the overall quality control procedure.

Data standardization

Global scaling was performed using the LogNormalize method, where each cell's total transcript count was normalized to 10,000 using a scaling factor (s0), then log-transformed for downstream analysis. CellCycleScoring was applied to determine cell cycle phases, and gene variability was assessed using the FindVariableFeatures function. To mitigate expression variability arising from mitochondrial and ribosomal transcript fractions and cell cycle differences, we performed regression correction with ScaleData. Linear dimensionality reduction of the expression matrix was achieved through RunPCA, after which significant principal components were selected for downstream analyses. Harmony was applied to correct batch effects, and UMAP was performed via RunUMAP for nonlinear dimensionality reduction.

Cell annotation

Cell type annotation was conducted by integrating information from the CellMarker and PanglaoDB repositories and relevant publications, further complemented by automated labeling via the SingleR tool. To identify marker genes for distinct cell populations, we employed the FindAllMarkers function, applying the criteria only. pos = TRUE and thresh. use = 0.25, meaning that exclusively upregulated markers expressed in no less than 25% of cells were retained.

Determine the contribution of cell subsets to the disease

The impact of distinct cellular subpopulations on disease was evaluated by analyzing shifts in their abundance alongside transcriptional dynamics. Initially, we determined unique marker genes for each subcluster. Briefly, differential expression analyses were carried out in batches to capture shifts in gene profiles between diseased and control samples. Next, we defined the FCscore: a tailored metric capturing simultaneous shifts in cell numbers and the expression levels of key genes participating in the biological process(Jin et al. 2022).

$$FC_{\mathrm{score}(i,j)} = \sqrt{FC_{\exp(i,j)} \cdot FC_{\mathrm{prop}(j)}} \quad (1)$$

The FCscore was designed as a heuristic composite metric to summarize relative transcriptional activity across cell populations rather than as a formal statistical estimator, following commonly adopted practices in single-cell analysis. In this context, $FC_{\exp(i,j)}$ represents the expression fold change of the *i*-marker within cluster *j*, whereas $FC_{\mathrm{score}(i,j)}$ indicates the relative proportion of fold change for the corresponding marker and cluster. Specifically, FCscore for marker *i* in cluster *j* is derived by normalizing $FC_{\exp(i,j)}$ with its square root. Ultimately, the influence of each cellular subset on disease is inferred from the mean FCscore computed across all marker genes in the respective cluster.

Functional analysis

Key genes were functionally annotated in Metascape (https://metascape.org), allowing detailed exploration of their biological context. Gene Ontology (GO) enrichment of the selected genes was performed with a minimum overlap of three hits and a significance cutoff of $p \leq 0.01$.

Mendelian randomization analysis

Robust MR analysis is based on three fundamental conditions: (1) Relevance — genetic instruments must exhibit a strong association with the exposure, without exerting a direct effect on the outcome; (2) Independence — instruments should not be correlated with any potential confounding variables; and (3) Exclusion — restriction, the effect of the instruments on the outcome must operate solely through the exposure of interest. If instrumental variables impact outcomes via alternative routes, this suggests the presence of genetic pleiotropy.

Outcome traits retrieved from the IEU OpenGWAS resource were drawn from GWAS summary statistics. Gene-expression associations in the eQTL dataset were harmonized by applying a locus-specific genome-wide threshold ($P < 1e\text{-}8$), and the resulting SNPs served as candidate instrumental variables (IVs). Pairwise linkage disequilibrium (LD) was then assessed, retaining variants with $r^2 < 0.001$ within a 10 Mb clumping window, followed by a secondary filter of $p_2 < 5 \times 10^{-8}$.

Causal inference was performed using several complementary approaches. These included inverse-variance weighting (IVW), which integrates SNP-specific Wald ratios via meta-analytic techniques; MR-Egger, which relaxes the assumption that instrument strength is independent of pleiotropic effects (InSIDE); the weighted median estimator, which remains consistent even when up to half of the instruments are invalid; and the weighted mode method, known for improving power, reducing bias, and lowering false-positive rates compared to MR-Egger. If only a single SNP was available for a causal estimate, the Wald ratio was solely applied.

These analytical approaches jointly yielded an integrated evaluation of how cis- and selected trans-acting gene regulation in peripheral blood contributes to bladder cancer susceptibility. Lastly, the validity of the identified causal associations was assessed using leave-one-out sensitivity analysis.

Sensitivity analysis

We applied leave-one-out sensitivity analysis within the MR framework to assess how individual genetic variants influence bladder cancer susceptibility. This approach identifies variants with disproportionate effects by sequentially excluding one SNP at a time and recalculating the combined effect estimate based on the remaining variants. Each exclusion produces a revised effect estimate accompanied by a 95% confidence interval, highlighting the impact of the omitted SNP and assessing the robustness of the overall results. The results present both the estimates obtained after removing individual SNPs and the overall estimate based on the full set of variants. By contrasting these values, we can determine if the exclusion of any single SNP significantly affects the combined outcome, thereby assessing the robustness of our findings.

Immune infiltration analysis

CIBERSORT is extensively employed to estimate the proportions of immune cell subsets within the tumor milieu. By applying support vector regression, it deconvolutes bulk gene expression data to distinguish immune populations. The method uses a signature matrix of 547 genes to differentiate 22 distinct immune cell phenotypes, including T cells, B cells, plasma cells, and diverse myeloid cell types. In the present study, CIBERSORT was employed on patient data to quantify the relative fractions of these immune cell categories. Subsequent Spearman correlation tests were conducted to explore the relationships between gene expression and immune infiltration levels.

GSEA analysis

Patients were classified into high- and low-expression groups according to key gene levels, followed by gene set enrichment analysis (GSEA) to identify differences in pathway activity between these subsets (Subramanian et al. 2005). Version 7.0 of the reference gene set was obtained from the MsigDB database. This curated collection of pathway subtypes facilitated comparative analysis between groups, allowing identification and ranking of gene sets showing significant enrichment with an adjusted p-value below 0.05. GSEA is frequently applied in studies aiming to connect disease stratification with underlying biological functions.

GSVA analysis

GSVA (Gene Set Variation Analysis) is an unsupervised method that employs nonparametric statistics to assess enrichment trends across sets of transcriptome-wide genes (Hänzelmann et al. 2013). Instead of examining changes at the individual gene level, GSVA captures alterations at the pathway scale by assigning enrichment scores to predefined gene collections, thereby revealing the functional characteristics of each sample. In this study, gene sets were sourced from the Molecular Signatures Database, and GSVA was utilized to calculate detailed scores for every set, with the goal of identifying functional variations across samples.

Transcription factor regulatory network

In this work, transcription factor identification was performed utilizing the RcisTarget R package, which relies exclusively on motif-based analyses. For each motif, the normalized enrichment score (NES) indicates the aggregate count of that motif within the reference database. Besides the motifs originally annotated in the datasets, we created supplementary annotation files by inferring motif similarities and aligning gene sequences. The enrichment of individual motifs within each gene collection was initially evaluated by calculating the area beneath the recovery curve (AUC) per motif -gene set combination, based on the comparison of gene set rankings to motif orderings. Subsequently, NES values for motifs were derived by situating each AUC within the distribution of all motif scores across the examined gene set.

Developmental trajectories of key cell subtypes

Research employing single-cell technologies enables the dissection of intricate physiological phenomena and the transcriptional dynamics within highly diverse cellular populations. Such analyses have uncovered genes that define specific cell subtypes, serve as indicators of intermediate biological states, or reflect transitions between distinct cellular fates. Frequently, such research reveals that gene expression varies asynchronously among individual cells, each capturing a unique temporal snapshot of the transcriptional landscape under study. Monocle provides an approach for arranging cells along pseudotime trajectories, leveraging this inherent asynchrony to map cells onto paths representative of biological processes like cellular differentiation.

Immunohistochemistry

To examine candidate gene expression at the tissue scale, immunohistochemical (IHC) staining was conducted on bladder cancer specimens preserved in formalin and embedded in paraffin. Archived radical cystectomy specimens were collected from Nanchong Central Hospital between March 2021 and March 2025, following approval by the institutional ethics committee. Eligibility required individuals to possess a confirmed primary malignancy of the bladder, to have undergone radical cystectomy, and to have accessible paired adjacent non-neoplastic tissue for examination.

Six antibodies (*ARHGEF18*, *HLA-DRB5*, *ISG20*, *NCF1*, *RPL13*, and *YPEL5*) were applied to assess protein expression in bladder cancer tissues and matched adjacent non-neoplastic counterparts. Immunoreactivity was defined by detectable staining in the membrane and/or cytoplasm. Protein expression levels were semi-quantitatively assessed based on staining intensity and the proportion of positive tumor cells, and samples were classified into low- and high-expression groups accordingly (detailed procedures and scoring criteria are described in the Supplementary Methods). All sections were independently evaluated by three senior pathologists blinded to clinicopathologic data, with final scores determined by consensus review.

Cell-based validation and in vitro functional assays

An immortalized normal urothelial cell line (SV-HUC-1) and two human bladder cancer cell lines (T24 and 5637) were employed for in vitro analyses. Basal expression levels of *ARHGEF18* and *YPEL5* were first compared between normal and malignant urothelial cells using RT–qPCR and western blotting. *ARHGEF18* silencing by siRNA and *YPEL5* overexpression via plasmid transfection were subsequently performed in T24 and 5637 cells, with modulation efficiency verified at both transcript and protein levels. The functional impact of gene modulation was then examined through cell viability/proliferation assays (CCK-8; 0, 24, 48, and 72 h), migration assays (wound healing; 0 and 48 h), invasion assays (Transwell; 48 h), and apoptosis analysis by flow cytometry (48 h). Detailed methodologies are described in the Supplementary Methods.

Statistical analysis

The Shapiro–Wilk method was applied to assess the distributional normality of continuous data. Variables following a Gaussian distribution are presented as mean ± standard deviation (SD) and were evaluated using parametric tests, such as paired or independent t-tests and one-way ANOVA followed by Tukey's multiple comparison procedure.

For variables failing normality assumptions, outcomes are presented as median and interquartile range (IQR) and evaluated using rank-based methods, including the Wilcoxon matched-pairs test, the Mann-Whitney U test, or the Kruskal-Wallis one-way ANOVA, as appropriate. The choice between Pearson's chi-squared test and Fisher's exact test was determined based on expected cell counts in contingency table analysis. Relationships among variables were assessed via Pearson or Spearman correlation metrics, depending on the underlying distribution.

Statistical analyses were carried out using R (v4.3.2) and GraphPad Prism (v10.0), with techniques selected according to data characteristics. A two-tailed P-value below 0.05 was regarded as statistically significant. Results were annotated as ns ($P \geq 0.05$), * ($P < 0.05$), ** ($P < 0.01$), ** ($P < 0.01$), *** ($P < 0.001$) and **** ($P < 0.0001$).

Results

Cellular subpopulations and functional profiles in bladder cancer

This investigation utilized transcriptomic data derived from 13 bladder cancer-associated tissue specimens. Initially, expression matrices were loaded through the Seurat toolkit, and cells were subsequently filtered according to total UMI levels, per-cell gene detection counts, and mitochondrial transcript percentages. Conventionally, observations deviating more than three times the median absolute deviation (MAD) from the central tendency were identified as extreme values and subsequently removed. Subsequently, the DoubletFinder tool was employed to identify and remove doublets, yielding a dataset comprising 72,664 high-quality cells, as depicted in violin and scatter plots (Supplementary Figure S1A-B). The ten genes exhibiting the greatest variability were illustrated (Supplementary Figure S1C).

Additionally, the optimal principal component (PC) count was determined via the ElbowPlot method, identifying 20 PCs (Supplementary Figure S1D). Principal component analysis (PCA) indicated evident inter-sample variability attributable to batch effects (Supplementary Figure S1E). To address this, Harmony integration was applied for further dimensionality reduction and batch correction (Supplementary Figure S1F), ultimately resolving the dataset into 16 distinct clusters through UMAP visualization (Fig. 1A). Each cluster underwent cell-type annotation, resulting in classification into 11 cell identities: Epithelial cells, $CD4^+$ T cells, Regulatory T cells, $CD8^+$ T cells, B cells, Plasma cells, Mast cells, Macrophages, Fibroblasts, Endothelial cells, and an unclassified group (Fig. 1B). A bubble plot visualizing classical markers for these 11 populations was produced (Fig. 1C), alongside a bar chart summarizing cell-type proportions within different groups (Fig. 1D).

Concurrently, the influence of individual cell subsets on bladder cancer was assessed by examining variations in both cell abundance and transcriptional profiles. Initially, differentially expressed genes distinguishing disease samples from controls were identified to characterize these shifts. A metric termed FCscore was defined to quantify combined changes in the expression magnitude and cell numbers of key genes implicated in biological pathways. This

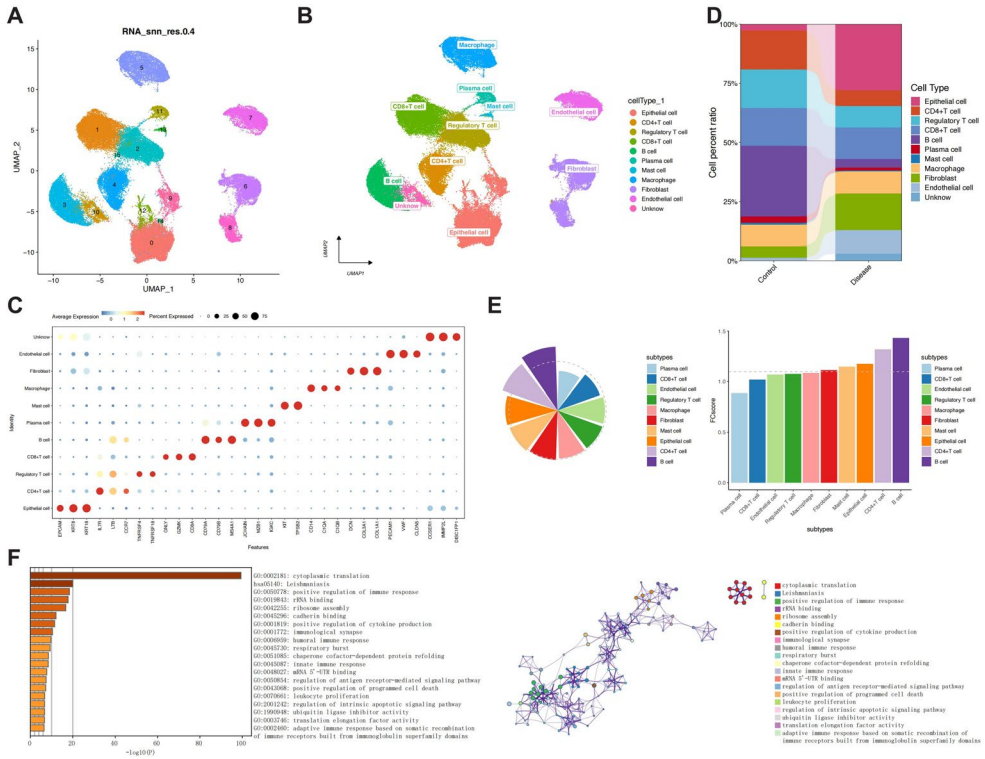

Fig. 1 Cellular Diversity in Bladder Cancer scRNA-seq: Comparative Profiling of Subtypes and Functional Interactions. **A** UMAP visualization depicting 16 distinct cellular clusters identified in bladder cancer tissues. **B** Subsequent cell-type annotation classified these clusters into 11 groups: Epithelial cells, $CD4^+$ T cell, regulatory T cell, $CD8^+$ T cell, B lymphocytes, Plasma cells, Mast cells, Macrophages, Fibroblasts, Endothelial cells, and a category designated as Unknown. **C** Bubble chart illustrating expression patterns of selected marker genes across cell populations, with the size of the dots reflecting the proportion of cells that express the marker and the color gradient denoting the average expression intensity. **D** Stacked bar graph showing relative proportions of various cell types in both control and disease cohorts. **E** Comparison of cell subset distribution and functional scoring (logFC = 0.585). **F** Pathway enrichment outcomes derived from analysis via the Metascape platform

analysis indicated that B cells contributed most prominently to bladder cancer pathology (Fig. 1E). The FindAllMarkers function was utilized to identify gene markers specific to each cell cluster (criteria: |logFC| > 0.585, adjPval < 0.05; detailed in Supplementary Table S1), with B cell-associated genes selected as key markers. Subsequently, pathway analysis of these marker genes was conducted via Metascape, uncovering marked enrichment in biological events, including upregulation of immune activity, Leishmaniasis-related pathways, and protein synthesis within the cytoplasm (Fig. 1F).

MR with multi-pronged validation identifies candidate genes driving bladder cancer

To further pinpoint crucial genes among the identified markers influencing bladder cancer, summary statistics encompassing 373,295 individuals (1,279 cases and 372,016 controls) were analyzed, corresponding to outcome ID ieu-b-4874. Instrumental variables and outcome data were sequentially retrieved through the use of extract_instruments() and extract_outcome_data() functions. Subsequent Causal inference between the six marker genes and eQTL-associated outcomes was conducted using MR methodology (Fig. 2A, IVW P value < 0.05) (refer to Supplementary Tables S2-8). The implicated genes were *ARHGEF18*, *HLA-DRB5*, *ISG20*, *NCF1*, *RPL13*, and *YPEL5*. Specifically, *ARHGEF18* (OR: 1.001; 95% CI: 1.000–1.002; $P = 0.017$), *HLA-DRB5* (OR: 1.000; 95% CI: 1.000–1.001; $P = 0.043$), *ISG20* (OR: 1.001; 95% CI: 1.000–1.001; $P = 0.029$), *NCF1* (OR: 1.001; 95% CI: 1.000–1.001; $P = 0.045$), and *RPL13* (OR: 1.001; 95% CI: 1.000–1.001; $P = 0.008$) were associated with heightened susceptibility to bladder malignancy, whereas *YPEL5* (OR: 0.997; 95% CI: 0.995–0.999; $P = 0.003$) exhibited a risk-reducing effect. Further robustness assessments were conducted to determine the reliability of the inferred causal associations. The findings demonstrated that omitting individual SNPs had minimal impact on overall confidence intervals, confirming the stability of the relationships observed for these six gene pairs (Fig. 2B). Accordingly, *ARHGEF18*, *HLA-DRB5*, *ISG20*, *NCF1*, *RPL13*, and *YPEL5* were identified as key targets for subsequent investigations.

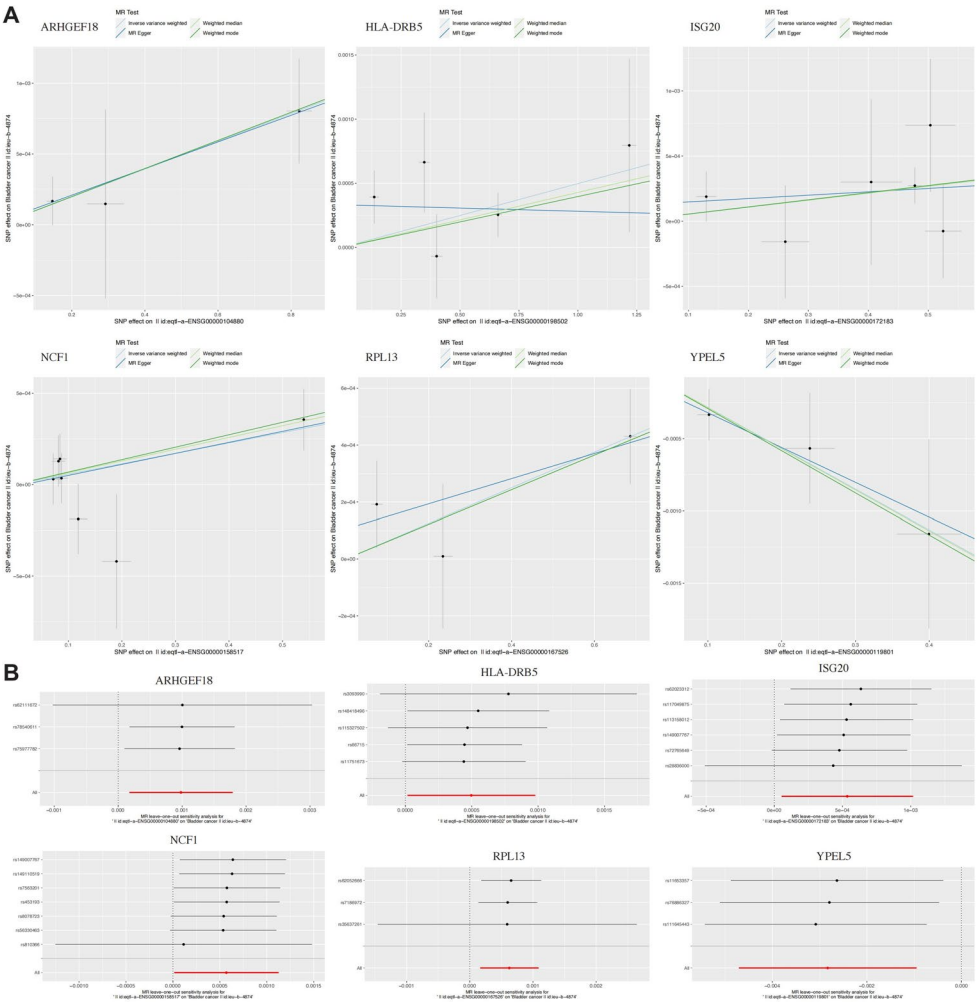

Fig. 2 Multi-Method MR Confirms Robust Causal Associations in Bladder cancer. **A** Scatter diagrams illustrating MR results for pivotal genes, namely *ARHGEF18*, *HLA-DRB5*, *ISG20*, *NCF1*, *RPL13*, and *YPEL5*. Distinct colors indicate various analytical methods, with line slopes reflecting the estimated causal effect specific to each approach. **B** Forest plots depicting leave-one-out analyses for SNPs linked to these key genes

Immune correlations of genes in tumor microenvironment

The tumor microenvironment consists primarily of stromal fibroblasts, various immune cell subsets, components of the extracellular scaffold, a spectrum of growth-promoting signals, pro-inflammatory agents, and distinct physicochemical conditions. This niche profoundly affects diagnostic accuracy, prognostic evaluation, and treatment responsiveness. Immune infiltration patterns and intercellular immune correlations were visualized through various representations (Fig. 3A-B). Compared with controls, samples from the disease condition group demonstrated notably higher infiltration of M1 macrophages and other immune subsets, whereas resting mast cells and several other cell types showed significant reductions (Fig. 3C).

Subsequent investigation examined the associations linking key genes with immune cell subsets. *ARHGEF18* exhibited a significant positive correlation with activated mast cells and resting NK cells, while showing an inverse relationship with resting mast cells. *HLA-DRB5* was positively associated with M1 macrophages and inversely related to activated dendritic cells. *ISG20* exhibited a negative correlation with activated dendritic cells. *NCF1* showed positive associations with both M1 and M2 macrophages. *YPEL5* showed a strong positive association with eosinophils (Fig. 3D).

Additionally, we investigated the relationships between key genes and a range of immune regulators, including stimulatory mediators, chemotactic cytokines, and their corresponding receptors. The results suggest that these genes are closely linked to immune infiltration patterns and may exert critical influence over the composition of the immune milieu (Fig. 3E).

Genes enrichment analysis and their connections to bladder cancer regulation and transcription factors

Subsequently, we explored the signaling routes linked to the identified genes to clarify possible molecular processes involved in disease progression. GSEA results revealed that *ARHGEF18* was primarily enriched in pathways related to ribosomal function, oxidative phosphorylation, and chemical carcinogenesis involving DNA adduct formation. *HLA-DRB5* was notably linked to chemokine activity, NF-κB signaling, and Toll-like receptor–mediated responses.

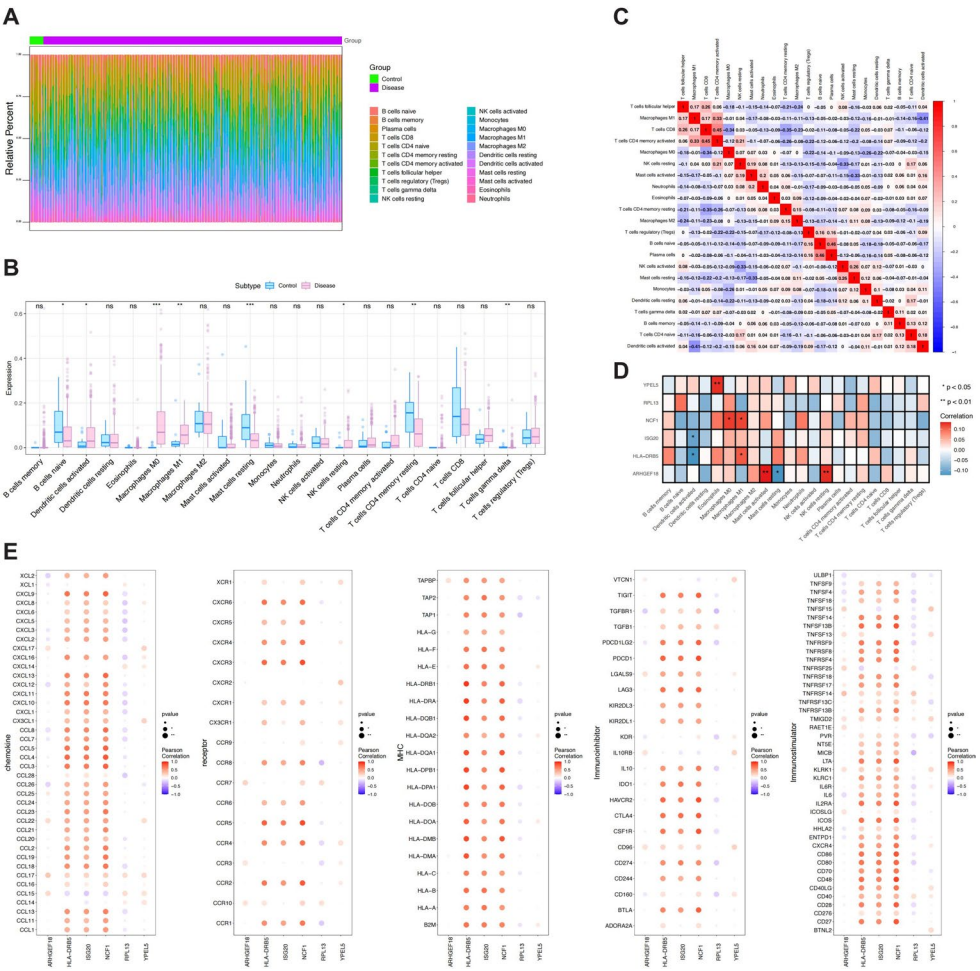

Fig. 3 Immunological Correlation Analysis of Key Genes. **A** Relative proportions of 22 distinct immune cell types in healthy controls versus disease samples. **B** Correlation matrix of the 22 immune subsets, where blue indicates inverse correlations and red denotes positive ones. **C** Differential distribution of immune cell populations across control and disease conditions; control samples are marked in blue, and disease samples in pink. **D** Correlative patterns between key genes and immune infiltration levels. **E** Relationships linking key genes with chemokines, immunosuppressive factors, immune-activating molecules, MHC components, and receptors

ISG20 showed involvement in NF-κB, chemokine, and IL-17 signaling pathways. Enrichment analysis of *NCF1* pointed to roles in chemokine, Toll-like receptor, and TNF signaling cascades. *RPL13* demonstrated marked enrichment in ribosomal biosynthesis, oxidative phosphorylation, and thermogenic processes. Lastly, *YPEL5* participated in cGMP–PKG, adipocytokine-mediated, and NF-κB signaling networks (Fig. 4A).

GSVA analysis (Fig. 4B and Supplementary Figure S2) indicated that *ARHGEF18* was predominantly enriched in pathways including xenobiotic metabolism, peroxisome processes, and others. *HLA-DRB5* exhibited enrichment in pathways related to allograft rejection, interferon gamma response, among additional signaling routes. Similarly, *ISG20* showed associations with allograft rejection, interferon gamma signaling, and similar pathways. *NCF1* was linked to allograft rejection, inflammatory processes, and other pertinent pathways. *RPL13* was notably enriched in oxidative phosphorylation, MYC targets v2, and other signaling networks. *YPEL5* demonstrated enrichment in early estrogen response, androgen response, and various other pathways. These observations imply that the highlighted genes could influence the advancement of the disease through the aforementioned pathways.

Genes implicated in disease modulation were retrieved via the GeneCards (https://www.genecards.org/) platform. To assess expression differences between groups, transcript levels of 20 high-relevance selected genes were analyzed. Significant variation was observed in *BRCA1*, *CDH1*, *PALB2*, *MSH2*, *MSH6*, *APC*, *BRIP1*, *BRCA2*, *TP53*, *CDKN2A*, *ATM*, *PTEN*, *POLE*, and *CHEK2* across patient subgroups (Fig. 5A). Furthermore, correlation analyses between six candidate genes and bladder cancer-associated regulators demonstrated substantial associations. Notably, *YPEL5* exhibited a positive association with *CDH1* (cor: 0.276, $P<0.05$), whereas *RPL13* showed a negative linkage to *PIK3CA* (cor: -0.533, $P<0.05$) (Fig. 5B).

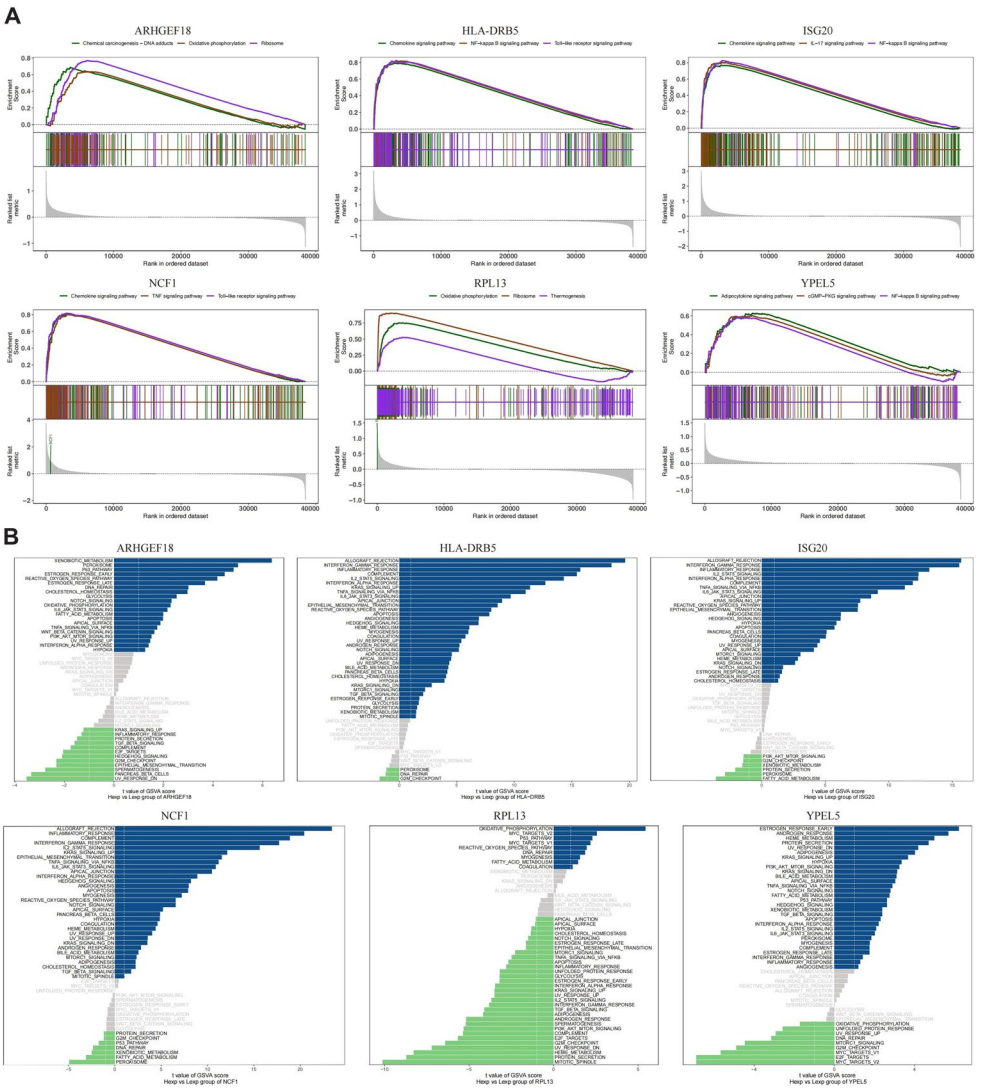

Fig. 4 GSEA and GSVA Insights for Key Genes. **A** KEGG pathways associated with key genes, illustrating pathway modulation and implicated gene participants. **B** GSVA outcomes for key genes. Blue signifies pathways linked to elevated gene expression, while green denotes pathways related to reduced gene expression. For background control, the Hallmark gene set was applied

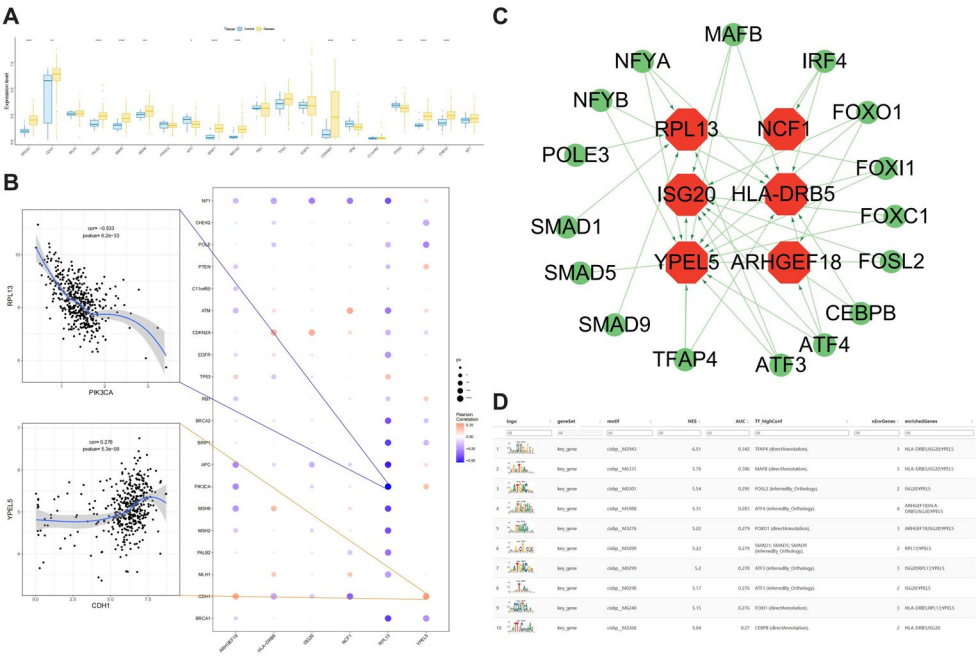

Fig. 5 Associations Between Key Genes, Disease Genes, and Transcriptional Regulation. **A** Transcript-level variations in disease-associated genes, where blue denotes the healthy cohort and pink corresponds to affected individuals. **B** An examination of the relationships between crucial genes and genes associated with diseases, with blue representing negative correlations and red denoting positive correlations. **C** Network map depicting transcriptional regulators of genes, where red nodes represent these key genes and green nodes denote the transcription factors. **D** Graphical representation of enriched motifs and their matched transcription factors linked to key genes

The identified signature genes served as input for downstream analysis, which revealed their potential regulation by common upstream elements, notably multiple transcriptional regulators. Accordingly, we conducted enrichment analysis of these transcription factors based on aggregate recovery curve evaluations. Motif-transcription factor pairing for pivotal genes highlighted cisbp_M2943 as the top-ranking motif, exhibiting the strongest normalized enrichment score (NES: 6.51). All overrepresented motifs linked to the key genes, along with their associated transcription factors, were visualized (Fig. 5C-D).

Role of key genes in single-cell data

Expression profiles of six key genes were analyzed based on single-cell transcriptomic data, demonstrating the distribution of *ARHGEF18*, *HLA-DRB5*, *ISG20*, *NCF1*, *RPL13*, and *YPEL5* across Epithelial cells, $CD4^+$ T cells, regulatory T cells, $CD8^+$ T cells, B lymphocytes, Plasma cells, Mast cells, Macrophages, Fibroblasts, Endothelial cells, and unclassified cell types (Fig. 6A-B). B lymphocytes were specifically selected for further investigation. Initially, cell-to-cell similarities were computed to establish differentiation trajectories. Visualization of these trajectories enabled the generation of pseudotime maps, capturing the dynamic progression of cell development and permitting analysis of differentiation processes alongside gene expression dynamics at various time intervals. Separate visual outputs were produced depicting cells colored by pseudotime scores (pseudotime denotes the probability estimated by Monocle2 using gene expression data to infer temporal ordering), cell identity, and cellular states defined by trajectory branch divisions (Fig. 6C-D). Genes are differentially expressed in different branches but cannot be displayed in the overall heat map. We selected all branch points, calculated the genes with large differences in cell expression before and after the branch points, and visualized them as branch heat maps. The starting point of the trajectory starts from the middle and differentiates to both sides, representing the differences at both ends of the branch point. The legend Cell Type's pre-branch is the data before the branch point, Cell fate1 is the side with a smaller state after the branch point, and Cell fate2 is the side with a larger state after the branch point. According to the changes in genes, it is divided into 6 clusters by default (Fig. 6E). Concurrently, variations in key gene expression across different cellular states and related features were illustrated (Fig. 6F).

AUCell was employed to measure activity scores of genes participating in immune metabolic pathways from the single-cell perspective, and differences in these activities for pivotal genes were visualized through bubble plots. Findings revealed that *YPEL5* exhibited pronounced activity within pathways such as MYC targets v1 and oxidative phosphorylation; *HLA-DRB5* was notably active in interferon gamma response and allograft rejection; *NCF1* showed elevated activity in allograft rejection and associated pathways; *ISG20* demonstrated higher activity across interferon alpha response, allograft rejection, and additional pathways; *RPL13* displayed increased activity in allograft rejection, interferon gamma response, and further pathways; while *ARHGEF18* presented heightened activity in MYC targets v1, oxidative phosphorylation, and other signaling routes (Fig. 6G).

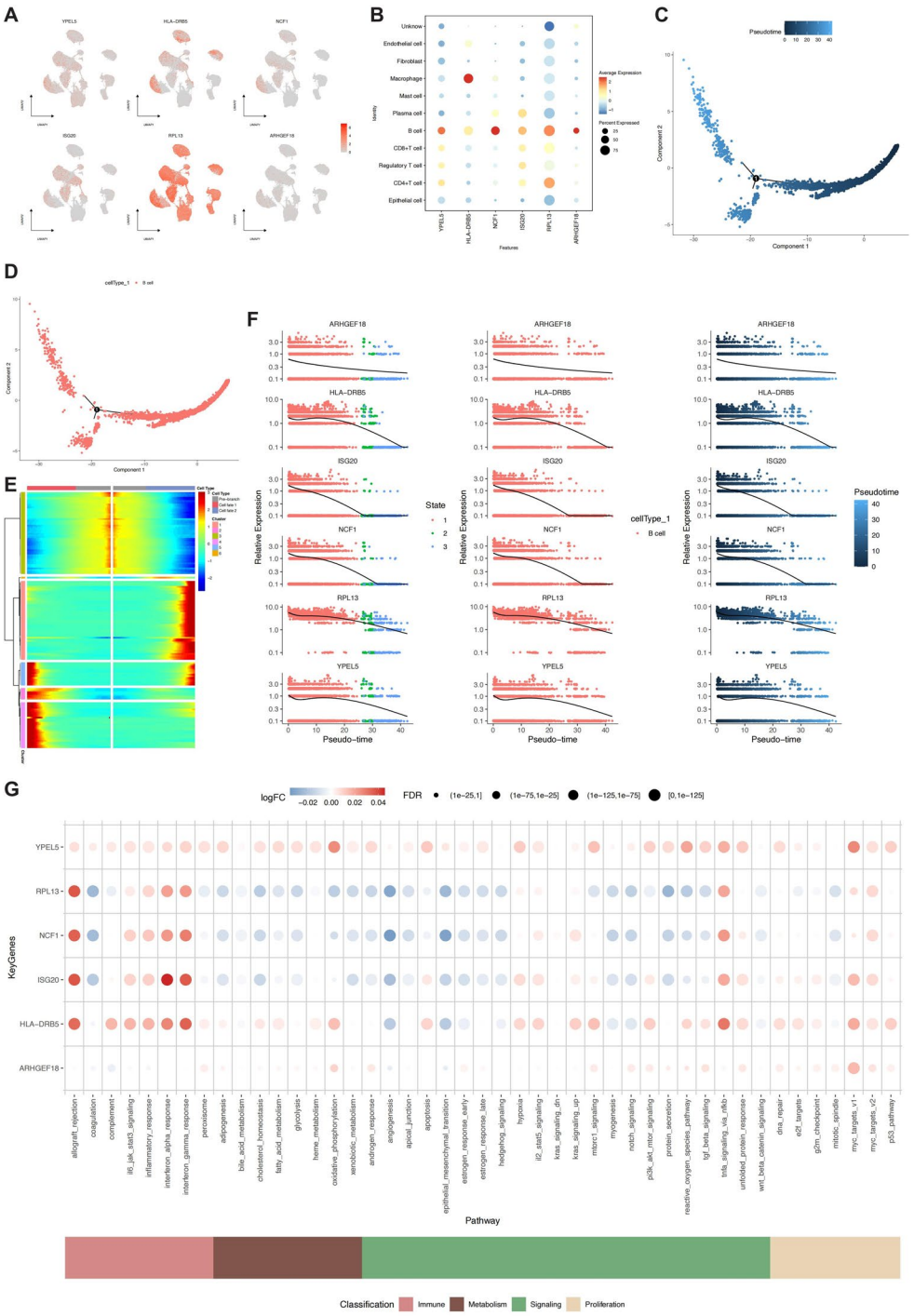

Fig. 6 Expression Patterns of Key Genes Across Cell Subtypes; Cellular Differentiation Paths and Metabolic Processes Modulated by Key Genes. **A** UMAP visualizations depicting distribution patterns of six pivotal genes (*YPEL5*, *HLA-DRB5*, *NCF1*, *ISG20*, *RPL13*, *ARHGEF18*) among various cellular clusters. **B** Bubble chart illustrating mean expression levels (color gradient) and proportions of cells expressing each gene (dot size) within distinct cell subtypes. **C**, **D** Temporal mapping of cellular states and differentiation trajectories. **E** Visualization of gene expression changes along branches of pigment cell lineages. **F** Associations between expression of key genes and cellular developmental pathways. **G** Bubble plot representing the relationship between genes and function enrichment, where blue reflects lower expression and red denotes higher expression

Gene expression differences and association with clinicopathological features

To explore the transcriptional profiles of candidate genes, four separate transcriptome datasets were examined. Within the TCGA-BLCA cohort, *YPEL5* exhibited markedly reduced expression in tumor samples relative to normal tissues, whereas *ARHGEF18* and *ISG20* were found to be elevated in malignancies. Notably, only *YPEL5* exhibited a statistically significant difference in paired sample analysis. *NCF1* was also downregulated, whereas *HLA-DRB5* and *RPL13* showed no significant differential expression (Fig. 7A). In GSE121711, similar trends were observed: *YPEL5* and *RPL13* were both downregulated, and *ARHGEF18* was upregulated in tumor tissues, with consistent results in paired comparisons (Fig. 7B). GSE13507 revealed lower expression of *YPEL5*, *HLA-DRB5*, *NCF1*, and *RPL13* in tumors, while GSE7476 confirmed downregulation of

YPEL5 and *HLA-DRB5*, and upregulation of *ARHGEF18*, though without paired samples (Fig. 7C-D).

Immunohistochemistry conducted on tumor specimens and matched adjacent normal tissues from 38 individuals with bladder cancer demonstrated uniform expression trends in (Supplementary Table S9 summarizes the scoring results) both paired and independent analyses. Specifically, *ARHGEF18* and *ISG20* showed notably elevated expression in tumor samples, whereas *HLA-DRB5* and *YPEL5* exhibited a significant decrease. These findings were supported by both continuous score comparisons and categorical analysis based on positive versus negative staining, further reinforcing the distinct expression profiles observed in tumor tissues in contrast with their adjacent normal counterparts (Fig. 7E-F). Molecular subtype–specific analyses were performed using the TCGA-BLCA cohort, which provides standardized and well-annotated subtype classifications, whereas the institutional cohort was primarily used for histological and protein-level validation.

Based on immunohistochemical staining results, patients were grouped into high and low expression categories for gene expression association analysis. The results revealed no significant correlation between gene expression and clinical pathological features (Supplementary Table S10-S15). This lack of correlation may be attributed to the limited sample size and missing follow-up data. Further analysis using the TCGA dataset produced similar results, showing a weak association between gene expression and clinical characteristics (Supplementary Table S16 and Supplementary Figure S3-S8).

Additionally, molecular subtype analysis based on gene expression in the TCGA cohort showed that in the high-expression group of *ARHGEF18*, the Luminal papillary subtype predominated, while the Basal squamous subtype was more common in the low-expression group. Similarly,

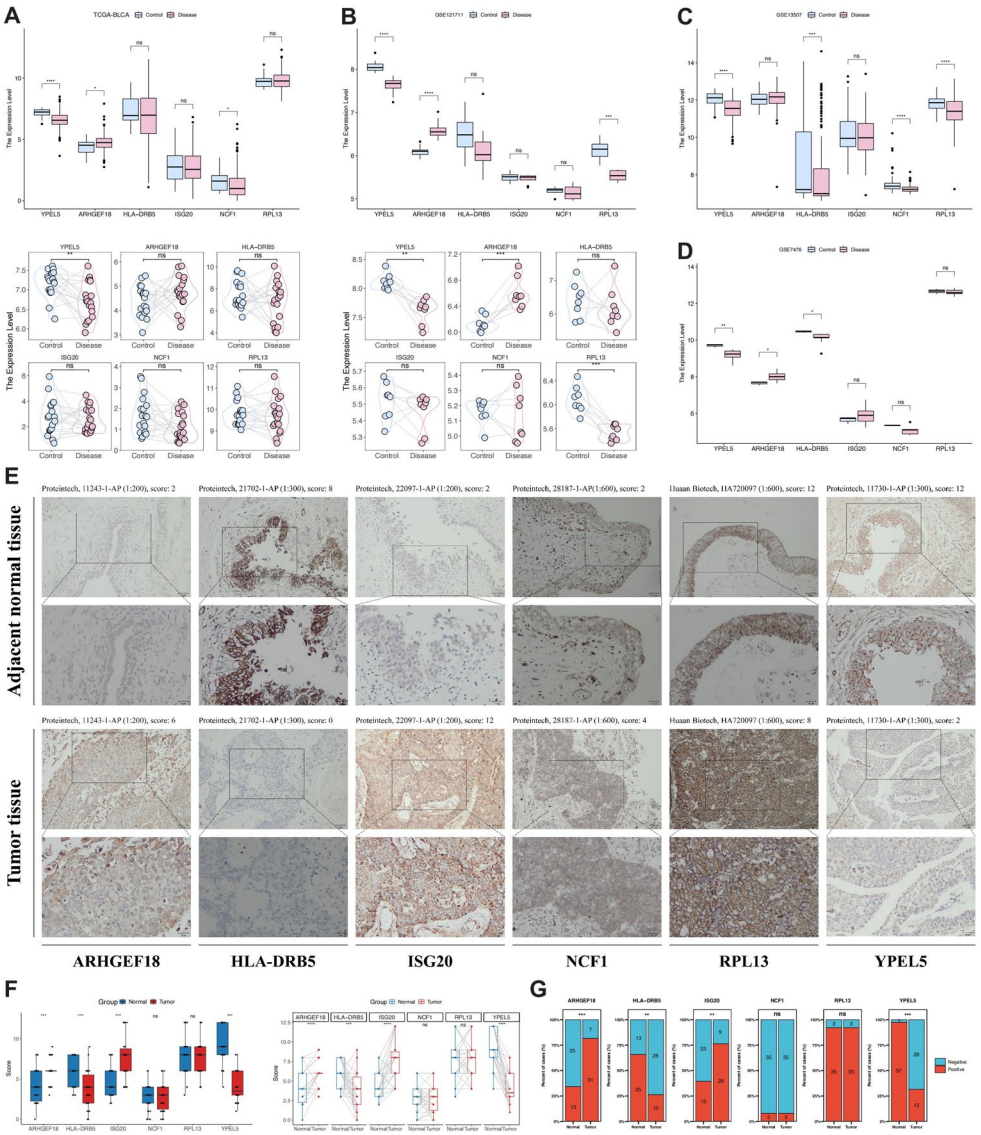

Fig. 7 Multi-cohort validation and immunohistochemical confirmation of candidate gene expression. **A**, **B** Expression differences of six candidate genes were assessed across TCGA-BLCA and GSE121711 datasets, covering unpaired as well as paired tumor-normal comparisons. **C**, **D** Expression profiles from GSE13507 and GSE7476 datasets using unpaired tumor and normal samples. **E** Representative immunohistochemical staining results for six genes in bladder cancer tissues and adjacent non-tumorous samples from 38 patients. **F** Quantitative immunohistochemical scoring comparisons between tumor and normal tissues using both unpaired and paired statistical analyses. **G** Summary of categorical analysis (positive vs. negative staining) for each gene in tumor and normal samples

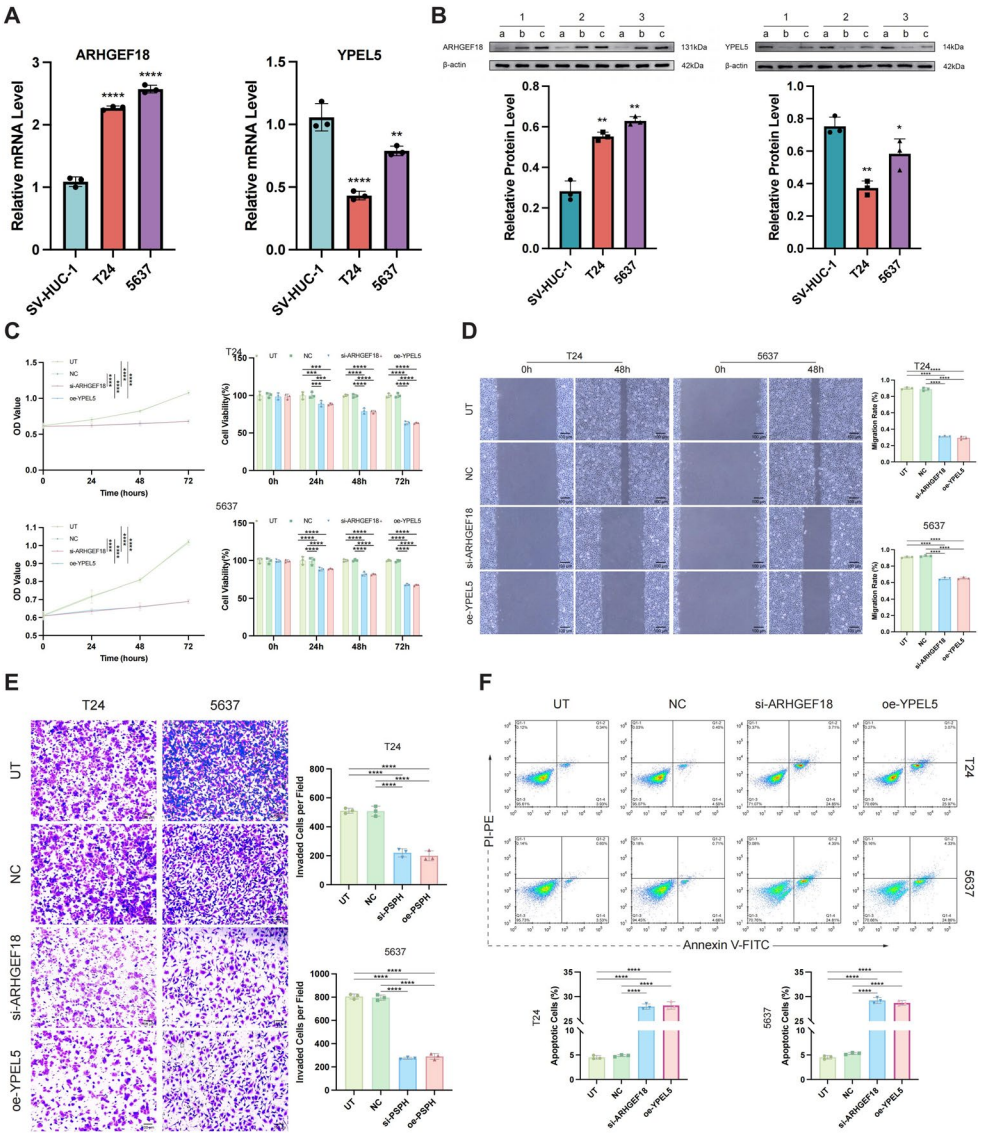

Fig. 8 In vitro analysis of *ARHGEF18* knockdown and *YPEL5* overexpression in bladder cancer cells. **A** qPCR analysis of *ARHGEF18* (left) and *YPEL5* (right) mRNA expression in T24 and 5637 bladder cancer cell lines, compared with normal urothelial cells (SV-HUC-1). Relative mRNA expression levels were quantified by qPCR. **B** Representative Western blot bands showing ARHGEF18 and YPEL5 protein levels in T24 and 5637 cells from three independent biological replicates. Densitometric quantification of protein levels is shown below the bands. **C** Cell proliferation assessed by CCK-8 assay in T24 (upper panel) and 5637 (lower panel) cells at 0, 24, 48, and 72 h. The left panels show OD values, and the right panels present relative proliferative activity at each time point. **D** Scratch wound-healing assay performed at 0 and 48 h. Representative images are shown in the left panels, and quantified migration rates at 48 h are presented in the right panels. **E** Transwell invasion assay at 48 h. Representative images of invaded cells are shown in the left panels, with quantitative analysis of invaded cell numbers shown in the right panels. **F** Flow cytometry analysis of apoptosis at 48 h. Representative flow cytometry plots are shown in the upper panels, and the percentages of apoptotic cells are summarized in the lower panels

for *YPEL5*, the Basal squamous subtype dominated in the low-expression group, while the Luminal papillary subtype was more prevalent in the high-expression group (Supplementary Figure S9). These findings emphasize the potential association between gene expression and specific molecular subtypes in bladder cancer. Based on the consistency and clarity of the expression patterns, *ARHGEF18* and *YPEL5* have been identified as key genes for further investigation in bladder cancer biology and prognosis.

In vitro analysis of *ARHGEF18* and *YPEL5* in bladder cancer cells

- Compared to normal urothelial cells (SC-HUC-1), the expression levels of *ARHGEF18* were significantly elevated in both T24 and 5637 cell lines, while the expression of *YPEL5* was comparatively lower. These findings were confirmed by qPCR and Western blot analyses (Fig. 8A-B). Subsequently, an *ARHGEF18* knockdown model was established in the T24 cell line using siRNA, with si-2747 exhibiting the most efficient silencing effect. Additionally, a plasmid for overexpressing *YPEL5* was successfully constructed, and its overexpression was validated at both the mRNA and protein levels through qPCR and Western blot (Supplementary Figure S10).

Phenotypic assays revealed significant alterations in cellular behavior following *ARHGEF18* knockdown and *YPEL5* overexpression. In the CCK-8 assay, there were no noticeable differences in OD values or proliferative activity at 0 h across all groups. However, over time, especially at 72 h, both the *ARHGEF18* knockdown and *YPEL5*

overexpression groups exhibited significantly reduced cell proliferation compared to the untreated (UT) and negative control (NC) groups (Fig.8 C). Scratch wound healing assays demonstrated a significantly lower healing rate in the knockdown and overexpression groups at 48 h, relative to the UT and NC groups (Fig.8D). Transwell invasion assays also showed a marked reduction in the number of invading cells in the knockdown and overexpression groups at 48 h compared to the UT and NC groups (Fig.8E). Flow cytometry analysis revealed a higher proportion of apoptotic cells in both the *ARHGEF18* knockdown and *YPEL5* overexpression groups compared to the UT and NC groups, further supporting the role of *ARHGEF18* and *YPEL5* in regulating cell proliferation, migration, and apoptosis in the T24 and 5637 cell lines (Fig.8 F).

Discussion

This study employed an integrative framework combining single-cell transcriptomics, Mendelian randomization, and experimental validation to identify candidate genes involved in bladder cancer. By linking cell-type–resolved transcriptional signals with genetically inferred causal effects, we aimed to move beyond descriptive expression patterns and provide biologically grounded evidence for disease-relevant regulators. Rather than introducing a new analytical framework, this study emphasizes biological interpretation and experimental validation of genetically supported candidate genes in bladder cancer. Through this approach, *ARHGEF18* and *YPEL5* were prioritized as key genes with consistent support across scRNA-seq analysis, genetic inference, bulk transcriptomic data, and functional assays. These findings highlight the value of integrating orthogonal data layers to improve the robustness of candidate gene identification in complex and heterogeneous malignancies such as bladder cancer (Teng et al. 2025; Wu et al. 2025).

Integrating genetically informed analysis with single-cell transcriptomic profiles offered a focused approach for prioritizing candidate genes within specific cellular contexts. Among all cell populations, B cells exhibited the most marked transcriptional alterations, particularly in pathways related to interferon signaling and antigen presentation (Yuen et al. 2016; Jiang et al. 2019; Laumont and Nelson 2023; Ma et al. 2024; Zhou et al. 2024). Based on these features, we identified differentially expressed genes within B cells and assessed their potential causal relevance to Bladder cancer using genetic instruments. This approach enabled us to move beyond descriptive associations and explore putative causal relationships (Chen et al. 2024a). Subsequent investigations using clinical specimens and cultured cell models further supported the functional involvement of selected genes. Together, this multi-level strategy improves biological interpretation while helping to mitigate spurious signals inherent to high-dimensional data (Teng et al. 2025).

Among the genes prioritized, *YPEL5* emerged as a particularly compelling tumor suppressor. It was consistently downregulated in both bladder cancer samples and cellular models, with MR analysis supporting its inverse association with disease risk. Functionally, *YPEL5* belongs to a conserved gene family involved in mitotic regulation; its knockdown prolongs G1, G2, and M phases, thereby impairing cell cycle transit and reducing proliferation (Hosono et al. 2004, 2010; Chen et al. 2018). Loss of *YPEL5* has also been shown to trigger mitochondrial membrane potential dissipation, implicating it in apoptosis regulation (Lee et al. 2017). Tumor-suppressive roles for *YPEL5* have been reported in cervical and colorectal cancers (Bokar et al. 1994; Barrett et al. 2013), and in colorectal cancer it is negatively regulated by the *METTL3-m6A-YTHDF2* signaling cascade, which facilitates tumor progression and metastatic potential (Zhou et al. 2021). Interestingly, the expression of *YPEL5* is increased in $CD4^+$ T cells of melanoma patients who respond to nivolumab, as well as in EGFR-mutant non-small cell lung cancer patients receiving erlotinib therapy (Wu 2018; Rad Pour et al. 2021), suggesting additional immune regulatory roles. In our dataset, *YPEL5* levels positively correlated with eosinophil infiltration, raising the possibility that it modulates anti-tumor immunity through cytokine or chemokine signaling.

In contrast, *ARHGEF18*, which encodes a Rho GTPase exchange factor, is notably upregulated in tumor specimens and enriched in oxidative phosphorylation and xenobiotic metabolism pathways (Kim et al. 2015; Arno et al. 2017). Rho GTPases act as key regulators of cytoskeletal dynamics, cellular polarity, and migration (Lin et al. 2010; Nakajima and Tanoue 2011; Struckhoff et al. 2011; Terry et al. 2012). *ARHGEF18* is critically involved in malignant tumor invasiveness, metastatic potential, and resistance to therapy through modulation of RhoA signaling and cytoskeletal remodeling, with its activity controlled by upstream factors including PKCγ and S100A10 (Zhao et al. 2021; Satow et al. 2022). Elevated *ARHGEF18* levels in lung squamous cell carcinoma correlate strongly with lymphatic metastasis and serve as a predictor of reduced overall survival (Song et al. 2013). Beyond its canonical roles in cytoskeletal remodeling, *ARHGEF18* has been identified via *CRISPR/Cas9* screening as a negative regulator of antibody secretion, suggesting a function in B-cell differentiation and humoral immunity (Trezise et al. 2023). It can also activate the *MAPK* pathway to induce Runx3 expression, affecting monocyte survival and differentiation (Dybska et al. 2023)—a mechanism potentially implicated in TAM development and the suppression of immune activity within the

tumor milieu (Dunsmore et al. 2024). Although its precise role in Bladder cancer requires further investigation, these findings support a model in which *ARHGEF18* promotes both tumor cell invasiveness and immune evasion.

This integrative approach offers several strengths. By overlaying genetically driven risk loci with cell-resolved transcriptomic data, we prioritized candidate genes that are not only statistically robust but also biologically contextualized. The combination of MR, scRNA-seq, and wet-lab validation helps filter out false positives common in high-dimensional analyses and strengthens biological interpretation (Quan et al. 2025; Wu et al. 2025). Furthermore, immune infiltration profiling and pseudotime trajectory analysis provide additional layers of functional insight, allowing us to infer how these genes may operate within dynamic tumor ecosystems rather than static cell populations (Liang et al. 2023; Yang et al. 2024). Taken together, this study demonstrates how integrating genetic causality with single-cell resolution can move beyond descriptive associations and enable biologically grounded prioritization of cancer drivers.

Several limitations of this study should be acknowledged. First, although the single-cell RNA sequencing dataset comprised a limited number of patient samples, it was intentionally used for exploratory, cell-type–resolved gene prioritization rather than for comprehensive quantification of tumor heterogeneity. Accordingly, conclusions were not drawn from rare or sparsely represented cell clusters, and the scRNA-seq analysis was primarily applied for hypothesis generation. To minimize potential bias, stringent quality control procedures were applied, yielding 72,664 high-quality cells, and the single-cell results were further supported by complementary analyses, including Mendelian randomization, bulk transcriptomic data from the TCGA-BLCA cohort, immunohistochemistry, and in vitro functional assays. Second, the Mendelian randomization analysis relied on blood-derived eQTL data, which may not fully reflect tissue-specific regulation in bladder cancer; however, the concordance observed across genetic, transcriptomic, and experimental evidence supports the biological relevance of the identified genes. Finally, the limited clinical sample size restricted detailed survival and subtype-stratified analyses, which will require validation in larger, well-annotated cohorts. Despite these limitations, the integrative analytical framework employed in this study provides a biologically grounded and scalable strategy for prioritizing candidate drivers in heterogeneous cancers.

Conclusion

This study combines scRNA-seq with genetic inference and experimental validation to uncover causal, cell-type-specific regulators in Bladder cancer. *YPEL5* and *ARHGEF18* emerge as robust candidates, illustrating a strategy for bridging genetic association and functional insight in heterogeneous tumor ecosystems.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00438-026-02371-w.

Acknowledgements We sincerely thank the contributors of the publicly available datasets used in this study, which provided a valuable foundation for our analyses. We are also deeply grateful to the patients and their families for generously providing surgical specimens. In addition, we appreciate everyone who contributed to this work in any capacity.

Author contributions Conceptualization, J.H. and J.W.; methodology, J.H., Z.J. and X.T.; software, S.P. and T.M.; validation, J.H., L.P. and Z.X.; formal analysis, J.H. and S.P.; investigation, Z.J. and L.P.; resources, Z.X. and S.P.; data curation, J.H.; writing—original draft preparation, J.H., Z.J., L.P. and T.M.; writing—review and editing, J.H., S.P., Z.X., X.T., Q.L. and J.W.; visualization, J.H. and Z.J.; supervision, J.W., Q.L. and Z.X.; project administration, J.W. and X.T.; funding acquisition, S.P., Z.X. and X.T. All authors have read and agreed to the published version of the manuscript.

Funding This study was supported by the University-level Fund of North Sichuan Medical College (CBY22-QNA48), the Nanchong Science and Technology Bureau (22SXQT0235), the Nanchong Government Science and Technology Program (22SXQT0233), and the Health Commission of Sichuan Province Medical Science and Technology Program (24QNMP084).

Data availability All data produced or examined throughout this research are contained within this article and its supplementary documents. Further details or original datasets can be obtained from the corresponding author upon a reasonable request.

Declarations

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

Ethical approval and consent to participate Approval from the Medical Ethics Committee of Nanchong Central Hospital was obtained for the pathological specimens used in this research, with the Approval Number: 2025 Audit (108). The ethics committee approved a waiver of informed consent, as the research posed minimal risk, did not affect patient rights or welfare, and could not be feasibly conducted without the waiver.

Consent for publication Not applicable.

References

Arno G, Carss KJ, Hull S, Zihni C, Robson AG, Fiorentino A et al (2017) Biallelic mutation of ARHGEF18, involved in the determination of epithelial apicobasal Polarity, causes Adult-Onset retinal degeneration. Am J Hum Genet 100:334–342. https://doi.org/10.1016/j.ajhg.2016.12.014

Bai Z, An J, Han J, Zhang Y, Wang H, Yang Z, Mo X (2025) Integrative genome-wide association studies, proteome-wide Mendelian randomization, and single-cell RNA sequencing identify interleukin-6 receptor protein as a therapeutic target in aortic aneurysm. Int J Biol Macromol 316:144380. https://doi.org/10.1016/j.ijbiomac.2025.144380

Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M et al (2013) NCBI GEO: archive for functional genomics data sets–update. Nucleic Acids Res 41:D991–995. https://doi.org/10.1093/nar/gks1193

Barrows D, Feng L, Carroll TS, Allis CD (2020) Loss of UTX/KDM6A and the activation of FGFR3 converge to regulate differentiation gene-expression programs in bladder cancer. Proc Natl Acad Sci U S A 117:25732–25741. https://doi.org/10.1073/pnas.2008017117

Bokar JA, Rath-Shambaugh ME, Ludwiczak R, Narayan P, Rottman F (1994) Characterization and partial purification of mRNA N6-adenosine methyltransferase from HeLa cell nuclei. Internal mRNA methylation requires a multisubunit complex. J Biol Chem 269:17697–17704

Boll LM, Vázquez Montes de Oca S, Camarena ME, Castelo R, Bellmunt J, Perera-Bel J, Albà MM (2025) Predicting immunotherapy response of advanced bladder cancer through a meta-analysis of six independent cohorts. Nat Commun 16:1213. https://doi.org/10.1038/s41467-025-56462-0

Chen M, Wei L, Law CT, Tsang FH, Shen J, Cheng CL et al (2018) RNA N6-methyladenosine methyltransferase-like 3 promotes liver cancer progression through YTHDF2-dependent posttranscriptional Silencing of SOCS2. Hepatology 67:2254–2270. https://doi.org/10.1002/hep.29683

Chen LG, Tubbs JD, Liu Z, Thach TQ, Sham PC (2024a) Mendelian randomization: causal inference leveraging genetic data. Psychol Med 54:1461–1474. https://doi.org/10.1017/s0033291724000321

Chen Y, Li E, Chang Z, Zhang T, Song Z, Wu H et al (2024b) Identifying potential therapeutic targets in lung adenocarcinoma: a multi-omics approach integrating bulk and single-cell RNA sequencing with Mendelian randomization. Front Pharmacol 15:1433147. https://doi.org/10.3389/fphar.2024.1433147

Dunsmore G, Guo W, Li Z, Bejarano DA, Pai R, Yang K et al (2024) Timing and location dictate monocyte fate and their transition to tumor-associated macrophages. Sci Immunol 9:eadk3981. https://doi.org/10.1126/sciimmunol.adk3981

Dybska E, Nowak JK, Walkowiak J (2023) Transcriptomic context of RUNX3 expression in monocytes: A Cross-Sectional analysis. Biomedicines

Essakly A, Loeser H, Kraemer M, Alakus H, Chon SH, Zander T et al (2020) PIK3CA and KRAS amplification in esophageal adenocarcinoma and their impact on the inflammatory tumor microenvironment and prognosis. Transl Oncol 13:157–164. https://doi.org/10.1016/j.tranon.2019.10.013

Gong M, Zhao X, Li Q, Hao Q, Cha L, Dong G et al (2025) Integrated analysis of bulk RNA sequencing, eQTL, GWAS, and Single-Cell RNA sequencing reveals key genes in hepatocellular carcinoma. J Cell Mol Med 29:e70359. https://doi.org/10.1111/jcmm.70359

Hänzelmann S, Castelo R, Guinney J (2013) GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics 14:7. https://doi.org/10.1186/1471-2105-14-7

Hosono K, Sasaki T, Minoshima S, Shimizu N (2004) Identification and characterization of a novel gene family YPEL in a wide spectrum of eukaryotic species. Gene 340:31–43. https://doi.org/10.1016/j.gene.2004.06.014

Hosono K, Noda S, Shimizu A, Nakanishi N, Ohtsubo M, Shimizu N, Minoshima S (2010) YPEL5 protein of the YPEL gene family is involved in the cell cycle progression by interacting with two distinct proteins RanBPM and RanBP10. Genomics 96:102–111. https://doi.org/10.1016/j.ygeno.2010.05.003

Hu X, Cai M, Xiao J, Wan X, Wang Z, Zhao H, Yang C (2024) Benchmarking Mendelian randomization methods for causal inference using genome-wide association study summary statistics. Am J Hum Genet 111:1717–1735. https://doi.org/10.1016/j.ajhg.2024.06.016

Jiang Q, Fu Q, Chang Y, Liu Z, Zhang J, Xu L et al (2019) CD19(+) tumor-infiltrating B-cells prime CD4(+) T-cell immunity and predict platinum-based chemotherapy efficacy in muscle-invasive bladder cancer. Cancer Immunol Immunother 68:45–56. https://doi.org/10.1007/s00262-018-2250-9

Jin K, Gao S, Yang P, Guo R, Li D, Zhang Y et al (2022) Single-Cell RNA sequencing reveals the Temporal diversity and dynamics of cardiac immunity after myocardial infarction. Small Methods 6:e2100752. https://doi.org/10.1002/smtd.202100752

Jubber I, Ong S, Bukavina L, Black PC, Compérat E, Kamat AM et al (2023) Epidemiology of bladder cancer in 2023: A systematic review of risk factors. Eur Urol 84:176–190. https://doi.org/10.1016/j.eururo.2023.03.029

Kim IH, Lee HJ (2020) Perioperative immunotherapy for muscle-invasive bladder cancer. Transl Androl Urol 9:2976–2985. https://doi.org/10.21037/tau.2019.11.31

Kim M, A MS, Ewald AJ, Werb Z, Mostov KE (2015) p114RhoGEF governs cell motility and lumen formation during tubulogenesis through a ROCK-myosin-II pathway. J Cell Sci 128:4317–4327. https://doi.org/10.1242/jcs.172361

Kim MC, Gate R, Lee DS, Tolopko A, Lu A, Gordon E et al (2024) Method of moments framework for differential expression analysis of single-cell RNA sequencing data. Cell 187:6393–6410e6316. https://doi.org/10.1016/j.cell.2024.09.044

Komura K, Hirosuna K, Tokushige S, Tsujino T, Nishimura K, Ishida M et al (2023) The impact of FGFR3 alterations on the tumor microenvironment and the efficacy of immune checkpoint inhibitors in bladder cancer. Mol Cancer 22:185. https://doi.org/10.1186/s12943-023-01897-6

Koutros S, Kiemeney LA, Pal Choudhury P, Milne RL, Lopez de Maturana E, Ye Y et al (2023) Genome-wide association study of bladder cancer reveals new biological and translational insights. Eur Urol 84:127–137. https://doi.org/10.1016/j.eururo.2023.04.020

Laumont CM, Nelson BH (2023) B cells in the tumor microenvironment: Multi-faceted organizers, regulators, and effectors of

anti-tumor immunity. Cancer Cell 41:466–489. https://doi.org/10.1016/j.ccell.2023.02.017
Lee JS, Leem SH, Lee SY, Kim SC, Park ES, Kim SB et al (2010) Expression signature of E2F1 and its associated genes predict superficial to invasive progression of bladder tumors. J Clin Oncol 28:2660–2667. https://doi.org/10.1200/jco.2009.25.0977
Lee JY, Jun DY, Park JE, Kwon GH, Kim JS, Kim YH (2017) Pro-Apoptotic role of the human YPEL5 gene identified by functional complementation of a yeast moh1Δ mutation. J Microbiol Biotechnol 27:633–643. https://doi.org/10.4014/jmb.1610.10045
Lerner SP, Tangen C, Svatek RS, Daneshmand S, Pohar KS, Skinner E et al (2024) Standard or extended lymphadenectomy for Muscle-Invasive bladder cancer. N Engl J Med 391:1206–1216. https://doi.org/10.1056/NEJMoa2401497
Liang T, Tao T, Wu K, Liu L, Xu W, Zhou D et al (2023) Cancer-Associated Fibroblast-Induced remodeling of tumor microenvironment in recurrent bladder cancer. Adv Sci (Weinh) 10:e2303230. https://doi.org/10.1002/advs.202303230
Lin KB, Tan P, Freeman SA, Lam M, McNagny KM, Gold MR (2010) The Rap GTPases regulate the migration, invasiveness and in vivo dissemination of B-cell lymphomas. Oncogene 29:608–615. https://doi.org/10.1038/onc.2009.345
Loras A, Suárez-Cabrera C, Martínez-Bisbal MC, Quintás G, Paramio JM, Martínez-Máñez R et al (2019) Integrative metabolomic and transcriptomic analysis for the study of bladder cancer. Cancers (Basel) 11. https://doi.org/10.3390/cancers11050686
Ma J, Wu Y, Ma L, Yang X, Zhang T, Song G et al (2024) A blueprint for tumor-infiltrating B cells across human cancers. Science 384:eadj4857. https://doi.org/10.1126/science.adj4857
Mengual L, Burset M, Ars E, Lozano JJ, Villavicencio H, Ribal MJ, Alcaraz A (2009) DNA microarray expression profiling of bladder cancer allows identification of noninvasive diagnostic markers. J Urol 182:741–748. https://doi.org/10.1016/j.juro.2009.03.084
Nakajima H, Tanoue T (2011) Lulu2 regulates the circumferential actomyosin tensile system in epithelial cells through p114RhoGEF. J Cell Biol 195:245–261. https://doi.org/10.1083/jcb.201104118
Patel VG, Oh WK, Galsky MD (2020) Treatment of muscle-invasive and advanced bladder cancer in 2020. CA Cancer J Clin 70:404–423. https://doi.org/10.3322/caac.21631
Quan W, Wang L, Xu J, Song J, Qin Y, Zeng H, Li J, Chen J (2025) Identifying key plasma proteins in the onset of parkinson's disease: proteome-wide mendelian randomization and single-cell RNA sequencing analysis. Mol Neurobiol 62:12153–12163. https://doi.org/10.1007/s12035-025-05041-x
Rad Pour S, Pico de Coaña Y, Demorentin XM, Melief J, Thimma M, Wolodarski M et al (2021) Predicting anti-PD-1 responders in malignant melanoma from the frequency of S100A9 + monocytes in the blood. J Immunother Cancer 9. https://doi.org/10.1136/jitc-2020-002171
Rafnar T, Vermeulen SH, Sulem P, Thorleifsson G, Aben KK, Witjes JA et al (2011) European genome-wide association study identifies SLC14A1 as a new urinary bladder cancer susceptibility gene. Hum Mol Genet 20:4268–4281. https://doi.org/10.1093/hmg/ddr303
Satow R, Suzuki Y, Asada S, Ota S, Idogawa M, Kubota S et al (2022) Downregulation of protein kinase C gamma reduces epithelial property and enhances malignant phenotypes in colorectal cancer cells. iScience 25:105501. https://doi.org/10.1016/j.isci.2022.105501
Song C, Gao Y, Tian Y, Han X, Chen Y, Tian DL (2013) Expression of p114RhoGEF predicts lymph node metastasis and poor survival of squamous-cell lung carcinoma patients. Tumour Biol 34:1925–1933. https://doi.org/10.1007/s13277-013-0737-8
Stein JP, Lieskovsky G, Cote R, Groshen S, Feng AC, Boyd S et al (2023) Radical cystectomy in the treatment of invasive bladder cancer: Long-Term results in 1,054 patients. J Clin Oncol 41:3772–3781. https://doi.org/10.1200/jco.22.02762
Struckhoff AP, Rana MK, Worthylake RA (2011) RhoA can lead the way in tumor cell invasion and metastasis. Front Biosci (Landmark Ed) 16:1915–1926. https://doi.org/10.2741/3830
Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA et al (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 102:15545–15550. https://doi.org/10.1073/pnas.0506580102
Teng F, Zhang R, Wang Y, Li Q, Wang B, Chen H et al (2025) Machine learning and Mendelian randomization reveal a tumor immune cell profile for predicting bladder cancer risk and immunotherapy outcomes. Am J Pathol 195:1141–1157. https://doi.org/10.1016/j.ajpath.2025.01.016
Terry SJ, Elbediwy A, Zihni C, Harris AR, Bailly M, Charras GT et al (2012) Stimulation of cortical myosin phosphorylation by p114RhoGEF drives cell migration and tumor cell invasion. PLoS ONE 7:e50188. https://doi.org/10.1371/journal.pone.0050188
Trezise S, Kong IY, Hawkins ED, Herold MJ, Willis SN, Nutt SL (2023) An arrayed CRISPR screen of primary B cells reveals the essential elements of the antibody secretion pathway. Front Immunol 14:1089243. https://doi.org/10.3389/fimmu.2023.1089243
van Hoogstraten LMC, Vrieling A, van der Heijden AG, Kogevinas M, Richters A, Kiemeney LA (2023) Global trends in the epidemiology of bladder cancer: challenges for public health and clinical practice. Nat Rev Clin Oncol 20:287–304. https://doi.org/10.1038/s41571-023-00744-3
Wang M, Li Z, Chu H, Lv Q, Ye D, Ding Q et al (2016) Genome-Wide association study of bladder cancer in a Chinese cohort reveals a new susceptibility locus at 5q12.3. Cancer Res 76:3277–3284. https://doi.org/10.1158/0008-5472.Can-15-2564
Wu X (2018) Up-regulation of YPEL1 and YPEL5 and down-regulation of ITGA2 in erlotinib-treated EGFR-mutant non-small cell lung cancer: A bioinformatic analysis. Gene 643:74–82. https://doi.org/10.1016/j.gene.2017.12.003
Wu X, Wang K, Li Q, Zhang Y, Wei P, Shan Y, Zhao G (2025) Combining Single-Cell RNA sequencing and Mendelian randomization to explore novel drug targets for parkinson's disease. Mol Neurobiol 62:7380–7392. https://doi.org/10.1007/s12035-025-04700-3
Yang E, Jing S, Wang F, Wang H, Fu S, Yang L et al (2024) Mesenchymal stem cells in tumor microenvironment: drivers of bladder cancer progression through mitochondrial dynamics and energy production. Cell Death Dis 15:688. https://doi.org/10.1038/s41419-024-07068-9
Yuen GJ, Demissie E, Pillai S (2016) B lymphocytes and cancer: a love-hate relationship. Trends Cancer 2:747–757. https://doi.org/10.1016/j.trecan.2016.10.010
Zhao Q, Zhang K, Li Z, Zhang H, Fu F, Fu J et al (2021) High migration and invasion ability of PGCCs and their daughter cells associated with the nuclear localization of S100A10 modified by sumoylation. Front Cell Dev Biol 9:696871. https://doi.org/10.3389/fcell.2021.696871
Zheng H, An M, Luo Y, Diao X, Zhong W, Pang M et al (2024) PDGFRα(+)ITGA11(+) fibroblasts foster early-stage cancer lymphovascular invasion and lymphatic metastasis via ITGA11-SELE interplay. Cancer Cell 42:682–700e612. https://doi.org/10.1016/j.ccell.2024.02.002
Zhou D, Tang W, Xu Y, Xu Y, Xu B, Fu S et al (2021) METTL3/YTHDF2 m6A axis accelerates colorectal carcinogenesis through epigenetically suppressing YPEL5. Mol Oncol 15:2172–2184. https://doi.org/10.1002/1878-0261.12898

Zhou J, Zhou R, Zhu Y, Deng S, Muhuitijiang B, Li C et al (2024) Investigating the impact of regulatory B cells and regulatory B cell-related genes on bladder cancer progression and immunotherapeutic sensitivity. J Exp Clin Cancer Res 43:101. https://doi.org/10.1186/s13046-024-03017-8

Zhu KL, Su F, Yang JR, Xiao RW, Wu RY, Cao MY et al (2024) TP53 to mediate immune escape in tumor microenvironment: an overview of the research progress. Mol Biol Rep 51:205. https://doi.org/10.1007/s11033-023-09097-7