INTERNATIONAL JOURNAL OF
CANCER & CELLULAR BIOLOGY RESEARCH
issn: 2641-4341
DOI: 10.37871

RESEARCH ARTICLE

Folate Metabolism-Associated Gene Signatures Reveal Novel Strategies for Prognosis and Immunotherapy Response in Bladder Cancer

Yangyang Man[1#], Biao Zhang[1], Yi Liu[1], Fei Yang[1], Chunhong Li[1], Jiahua Hu[2] and Xianliang Hou[1,2*]

[1]Laboratory Center, Guangxi Key Laboratory of Metabolic Reprogramming and Intelligent Medical Engineering for Chronic Diseases, the Second Affiliated Hospital of Guilin Medical University, Guilin, 541199, China
[2]Scientific Research Department, the Second Affiliated Hospital of Guilin Medical University, Guilin, 541199, China
[#]Equally Contributed

Abstract

The mechanisms through which folate metabolism influences Bladder Cancer (BLCA) prognosis and the tumor immune microenvironment remain insufficiently understood. By integrating multi-omics data, we identified 341 genes associated with Folate Metabolism (FAMGs) and developed a 6-gene Folate-Metabolism-Derived Prognostic Index (FAMPI: SLC19A3, MTHFD1L, CAV1, POU5F1, SETBP1, and HSPG2) using machine learning. In both the TCGA-BLCA training set and the independent GSE13507 cohort, elevated FAMPI scores were consistently linked to poorer overall survival, whereas lower scores suggested better outcomes. A nomogram combining FAMPI and clinical-pathological variables surpassed individual clinical features and offered the highest net clinical benefit in decision-curve analysis. On a biological level, high-FAMPI tumors exhibited an inflammatory but immunosuppressive microenvironment, characterized by elevated checkpoint molecule expression (e.g., CD274 and CTLA4), increased TIDE and T-cell exclusion scores, and enhanced immune-cell infiltration—indicating immune activity alongside immune evasion and dysfunction. In contrast, the Immunophenoscore (IPS) was higher in the low-FAMPI group, implying a better response to immune-checkpoint inhibitors. Drug-response modeling also demonstrated distinct sensitivities to small-molecule/targeted inhibitors (e.g., IGF1R and HSP90 inhibitors, dasatinib) between the FAMPI risk groups, supporting the use of FAMPI in guiding treatment choices. Overall, FAMPI is a reliable prognostic classifier and a potential indicator for forecasting treatment outcomes in BLCA, offering molecular insights for personalized strategies for treatment.

***Corresponding author(s)**

Xianliang Hou, Laboratory Center, Guangxi Key Laboratory of Metabolic Reprogramming and Intelligent Medical Engineering for Chronic Diseases, the Second Affiliated Hospital of Guilin Medical University, Guilin, 541199, China

Email: houxl115@126.com

DOI: 10.37871/ijccbr24

Submitted: 06 January 2026

Accepted: 28 January 2026

Published: 31 January 2026

Keywords

- Bladder cancer; Folate metabolism; Machine learning; Immunotherapy; Drug sensitivity

VOLUME: 5 ISSUE: 1

How to cite this article: Man Y, Zhang B, Liu Y, Yang F, Li C, Hu J , Hou X. Folate Metabolism-Associated Gene Signatures Reveal Novel Strategies for Prognosis and Immunotherapy Response in Bladder Cancer. Int J Cancer Cell Biol Res. 2026 Jan 31; 5: 22. doi: 10.37871/ijccbr24

INTRODUCTION

Bladder Cancer (BLCA) is among the most common malignancies of the urinary system and is widely recognized for its strong tendency toward recurrence and metastasis. Its global incidence and mortality remain elevated, with a significantly greater prevalence in males than in females [1,2]. Urothelial carcinoma, squamous cell carcinoma, and adenocarcinoma make up the majority of BLCA, with Urothelial Bladder Cancer (UBC) being the most frequent, representing more than 90% of cases [3]. Major risk factors for BLCA include cigarette smoking, chronic exposure to chemical carcinogens, and chronic cystitis [4]. Although early-stage BLCA can be diagnosed by urinalysis and cystoscopy and treated with surgery, chemotherapy, or radiotherapy [5], some patients are already diagnosed with distant metastasis. In addition, postoperative recurrence is common, and the efficacy of existing second-line therapies is limited, making advanced BLCA particularly challenging to treat [6]. Immunotherapy, and more specifically Immune Checkpoint Inhibitors (ICIs), has seen major developments in managing BLCA in recent years [7]. However, tumor heterogeneity and differences in the immune microenvironment lead to considerable variation in patient responses, with overall response rates remaining unsatisfactory [8]. These limitations highlight the urgent need to explore metabolic pathways such as folate metabolism, which may shape both tumor biology and the immune microenvironment. Hence, it is essential to pinpoint unique and cutting-edge molecular biomarkers and to establish effective predictive models to guide personalized treatment.

Folate metabolism, a central component of one-carbon metabolism, is essential for nucleotide synthesis, S-Adenosylmethionine (SAM) generation, and DNA methylation [9]. In normal cells, folate metabolism maintains a dynamic balance between nucleic acid synthesis and epigenetic regulation. However, in tumor cells, this pathway is frequently reprogrammed due to genetic mutations, enzymatic abnormalities, or altered nutrient availability, leading to accelerated DNA synthesis, aberrant epigenetic modifications, and genomic instability [10]. Accumulating evidence indicates that aberrant expression related to folate metabolism is intimately linked to the initiation and the advancement of different forms of cancer. Moreover, these genes may modulate the immune microenvironment of the tumor by adjusting immune cell infiltration and checkpoint molecule expression, there by affecting responsiveness to immunotherapy [11]. For example, high expression of one-carbon/folate metabolism-related genes is linked to altered immune infiltration and unfavorable outcomes in breast cancer [12]. Similarly, dysregulation of FAMGs in colorectal cancer has been linked to immunosuppressive microenvironments and adverse clinical outcomes [13], suggesting that folate metabolism may modulate tumor progression through a metabolism–immunity axis. However, systematic investigations into the prognostic significance, immunological associations, and therapeutic potential of FAMGs in BLCA remain limited.

To address this deficiency, we synthesized transcriptomic and clinical data from TCGA and GEO cohorts and applied machine learning strategies to formulate a Folate Metabolism–Based Prognostic Index (FAMPI). Using FAMPI scores, we conducted Patient stratification according to FAMPI scores identified distinct molecular and clinical subtypes, providing new perspectives on the involvement of folate metabolism in BLCA progression and immune modulation. This research therefore introduces an innovative molecular tool for individualized therapy and contributes to a deeper mechanistic understanding of folate metabolism in bladder cancer.

MATERIALS AND METHODS

Dataset collection

In this investigation, transcriptomic sequencing data along with associated clinical details for Bladder Cancer (BLCA), encompassing both tumor and normal control samples, were

obtained from the TCGA database (https://portal.gdc.cancer.gov/) [14]. The raw sequencing data were normalized, and tumor samples labeled as "01A" and normal samples labeled as "11A" were saved for subsequent analysis. An expression matrix was then constructed employing protein-coding genes, with tumor sample expression normalized against normal samples. This matrix was subsequently used as the basis for analyzing differential expression [15,16]. and model construction.

Differential expression analysis

Expression data normalization was performed using the voom function [15]. We employed the R package called limma [16] to investigate the Differential Expression Of Folate Metabolism-Associated Genes (DEFAMGs) between tumor (01A) and normal (11A) samples. Genes were considered differentially expressedif they satisfied the conditions of |log2FC| being greater than 1 and a p value less than 0.05 [16]. To establish a Protein-Protein Interaction (PPI) network, significantly altered genes were uploaded to the STRING database [17]; the resulting network was subsequently visualized through Cytoscape (version 3.8.2) [18] Using the MCODE plugin [19] in Cytoscape, key modules were identified with the following parameters: degree cutoff of 2, node score cutoff of 0.2, k-core of 2, and a maximum depth of 100. Hub DEFAMGs were subsequently identified employing five different algorithms such as stress, betweenness, radiality, closeness, and bottleneck within the "CytoHubba" plugin [20]. Finally, using NetworkAnalyst 3.0, predictions were made for the transcription factors of the significant genes [21] to construct a transcription factor–target gene regulatory network.

Identification of famgs-associated clusters

Consensus clustering was performed on the expression profiles exhibiting Differential Expression Folate Metabolism-Related Genes (DEFAMGs) using the R package Consensus Cluster Plus [22]. Evaluation of DEFAMG expression levels within the TCGA-BLCA dataset identified two clusters as optimal, dividing all tumor samples into subtypes C1 and C2. Survival differences between these subtypes were subsequently visualized employing the R survival package [23].

Weighted gene co-expression network analysis

To examine folate metabolism-related genes with potential key functions in critical modules, the R package WGCNA was employed to develop a weighted network of the coordinated expression of multiple genes [24].Based on the normalized expression matrix, low-expression genes and outlier samples were first removed. A soft approach was used to build a scale-free network threshold approach power selection, after which genes were clustered according to the Topological Overlap Matrix (TOM) and assigned to distinct co-expression modules. Correlations between Module Eigengenes (MEs) and sample subtypes (C1/C2) were further analyzed to identify key modules that were strongly associated with specific subtypes. We also examined the associations between MEs, tumor stage, and overall survival, and found that the turquoise module showed the strongest correlation with the folate-metabolism–related subtype (cluster2), more advanced pathological stage, and poorer prognosis. Therefore, the turquoise module was defined as the hub module for subsequent modeling and mechanistic studies.Development of prognostic signature

To construct a folate metabolism–related prognostic signature, we employed an integrated multi-algorithm machine learning pipeline. Genes in the turquoise module identified by WGCNA were intersected with DEFAMGs, yielding 74 candidate genes. Subsequently, univariate Cox proportional hazards regression was applied to these 74 genes with a significance threshold of $p < 0.05$ to obtain candidate prognostic FAMGs. These candidate genes were then independently evaluated by four machine learning algorithms: LASSO Cox regression [25], XGBoost [26], decision tree

[27], and random forest [28]. Each algorithm produced a gene importance ranking, with LASSO using the absolute values of regression coefficients and the tree-based models using feature importance/Gini importance. The penalty parameter λ in the LASSO Cox model was determined ***via*** 10-fold cross-validation. To integrate the results of different algorithms and reduce bias from any single method, we adopted a consensus-ranking strategy: the rankings of each gene across the four algorithms were min–max normalized and then averaged with equal weights to obtain a consensus importance score, from which the top 10 genes were selected as the core candidate set for constructing the prognostic model. These 10 consensus genes were subsequently entered into a multivariate Cox proportional hazards model, and stepwise regression was used for variable selection. This process ultimately yielded an optimal six-gene signature, termed the Folate Metabolism–Associated Prognostic Index (FAMPI), consisting of SLC19A3, MTHFD1L, CAV1, POU5F1, SETBP1, and HSPG2. For each patient, the FAMPI risk score was calculated as follows: FAMPI score = ($coefficient_1$ × $expression\ value_1$) + ($coefficient_2$ × $expression\ value_2$) + … + ($coefficient_6$ × $expression\ value_6$), where $coefficient_1$–$coefficient_6$ represent the multivariate Cox regression coefficients for SLC19A3, MTHFD1L, CAV1, POU5F1, SETBP1, and HSPG2, respectively, and $expression\ value_1$–$expression\ value_6$ denote the normalized expression values of these genes.

Development of the nomogram

We employed multivariate Cox regression and applying stepwise regression analyses to combine characteristics of patients with BLCA, including age, gender, clinical pathological stage, and FAMPI, to create a predictive nomogram [29]. Visualizations of the nomogram and calibration plot were generated using the rms package [30]. The time ROC package was utilized to perform a Receiver Operating Characteristic (ROC) analysis for patients with BLCA [31]. FAMPI was analyzed through correlation and stratification based on the specified clinical parameters. Furthermore, Decision Curve Analysis (DCA) was used to assess the net benefit of combining the nomogram with a model that relies solely on clinical characteristics [32].

Gene set enrichment analysis

Potential molecular mechanisms distinguishing the high- and low-risk groups, as determined by the FAMPI risk model, were examined through Gene Set Enrichment Analysis (GSEA) conducted with the GSEA software (version 4.2.3) [33]. Gene sets from hallmark, KEGG, and GO in MSigDB were used as references [34-37]. Standard parameters were applied according to the GSEA guidelines, and significance thresholds were set at FDR < 0.25 and NOM $p < 0.05$ to identify enriched pathways.

Tumor immune microenvironment analysis

To evaluate the degree of immune cell presence in BLCA patients, the tumor immune microenvironment was analyzed using ESTIMATE, ssGSEA, and CIBERSORT. The immune score, stromal score, and tumor purity corresponding to each sample were quantified through the ESTIMATE algorithm [38]. The comparative abundance of 28 immune cell types was gauged ***via*** ssGSEA [39]. And the CIBERSORT algorithm was harnessed to quantify the fractions of 22 immune cell subsets in each sample [40]. To explore potential mechanisms in the immune microenvironment, disparities in immune cell infiltration between high- and low-risk groups delineated by the FAMPI risk model were investigated.

Mutation profile analysis

Variations in mutation profiles across the FAMPI model–defined high- and low-risk groups were analyzed using mutation data from the TCGA-BLCA cohort [41], processed with the R package maftools [42]. Mutation frequency and waterfall plots were subsequently generated for each group. We further compared Tumor Mutation Burden (TMB) between the two groups [43] to assess mutation level distributions across subtypes, thereby elucidating potential

molecular mechanisms associated with the model.

Identification of the immunotherapy efficacy

The Cancer Immunome Atlas (TCIA, https://tcia.at/home) provided data to gauge how BLCA patients might respond to Immune Checkpoint Inhibitor (ICI) therapy [44]. The TIDE algorithm was employed to simulate anti–PD-1 and anti–CTLA-4 therapeutic scenarios, estimating differences in immunotherapy responses between FAMPI model-defined high- and low-risk groups, thereby providing additional insights into immune evasion capacity and immunotherapeutic sensitivity [45].

Drug sensitivity analysis

The oncoPredict package in R was developed to quantify the chemosensitivity of TCGA-BLCA patients stratified by FAMPI risk scores, aiming to facilitate individualized therapeutic strategies [46]. To determine the concentration at which inhibition is half-maximal (IC50), the oncoPredict package was utilized by aligning data on gene expression from patient samples with that of cancer cell lines. The Wilcoxon rank-sum test was applied to assess differences in predicted IC50 values between high- and low-risk groups, with a importance threshold of $p < 0.05$ [47]. To enhance the reliability of drug sensitivity evaluation, the GSCALite platform was additionally utilized, which integrates the GDSC, CTRP, and CellMiner datasets (http://bioinfo.life.hust.edu.cn/GSCA/#/) [48]. Notable variations in predicted drug sensitivity were observed between groups classified as high-risk and low-risk by the FAMPI model. Statistical analysis

Survival curves were generated through the Kaplan–Meier method [29]. The Wilcoxon test [49] was employed to compare two groups, whereas the Kruskal–Wallis test was utilized for comparisons among multiple groups. The relationships between variables were assessed using Spearman's rank correlation analysis [50]. A P value threshold of 0.05 or less defined statistical significance, and all analyses were performed using R software (version 4.1.3).

RESULTS

Identification of differentially expressed DEFAMGs-related genes

Through comparison of Tissues from BLCA and healthy tissues, we discovered 4,486 Differentially Expressed Genes (DEGs), of which 2,120 were up-regulated and 2,366 genes were down-regulated in BLCA patients. Intersecting these DEGs with Folate Metabolism–Associated Genes (FAMGs) yielded 341 overlapping genes, which were defined as DEFAMGs. The volcano plot of DEFAMGs and the Venn diagram illustrating the overlap between DEGs and FAMGs are shown in figures 1A,B. The overlapping genes between BLCA-DEGs and FAMGs were subjected to PPI network construction using the STRING database and subsequently illustrated in Cytoscape (v3.8.2), resulting in the interaction network shown in figure 1C. Eight overlapping hub genes (FGF2, FGF7, FGF10, IL6, IGF1, VCAM1, STAT3, and CXCL12) were identified by employing five different algorithms (MCC, EPC, Degree, Betweenness, and Closeness) in Cytoscape (v3.8.2) (Figures 1D,E). Gene Ontology (GO) enrichment analysis revealed that the DEFAMGs were primarily related to cellular responses to inorganic substances, lipid metabolism, xenobiotic stimuli, and cell proliferation regulation and phosphorylation (Figures 1F,G). The marked enrichment of DEFAMGs in multiple cancer-related pathways—including Pathways in cancer, PI3K–Akt signaling, Fanconi anemia, and MicroRNAs in cancer—was identified through Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis (Figure 1H).

Identification of Clusters related to FAMG-Related Genes

To stratify the BLCA cohort, we employed consensus clustering on the expression patterns of the eight overlapping hub DEFAMGs, evaluating cluster numbers (k) from 2 to 8.

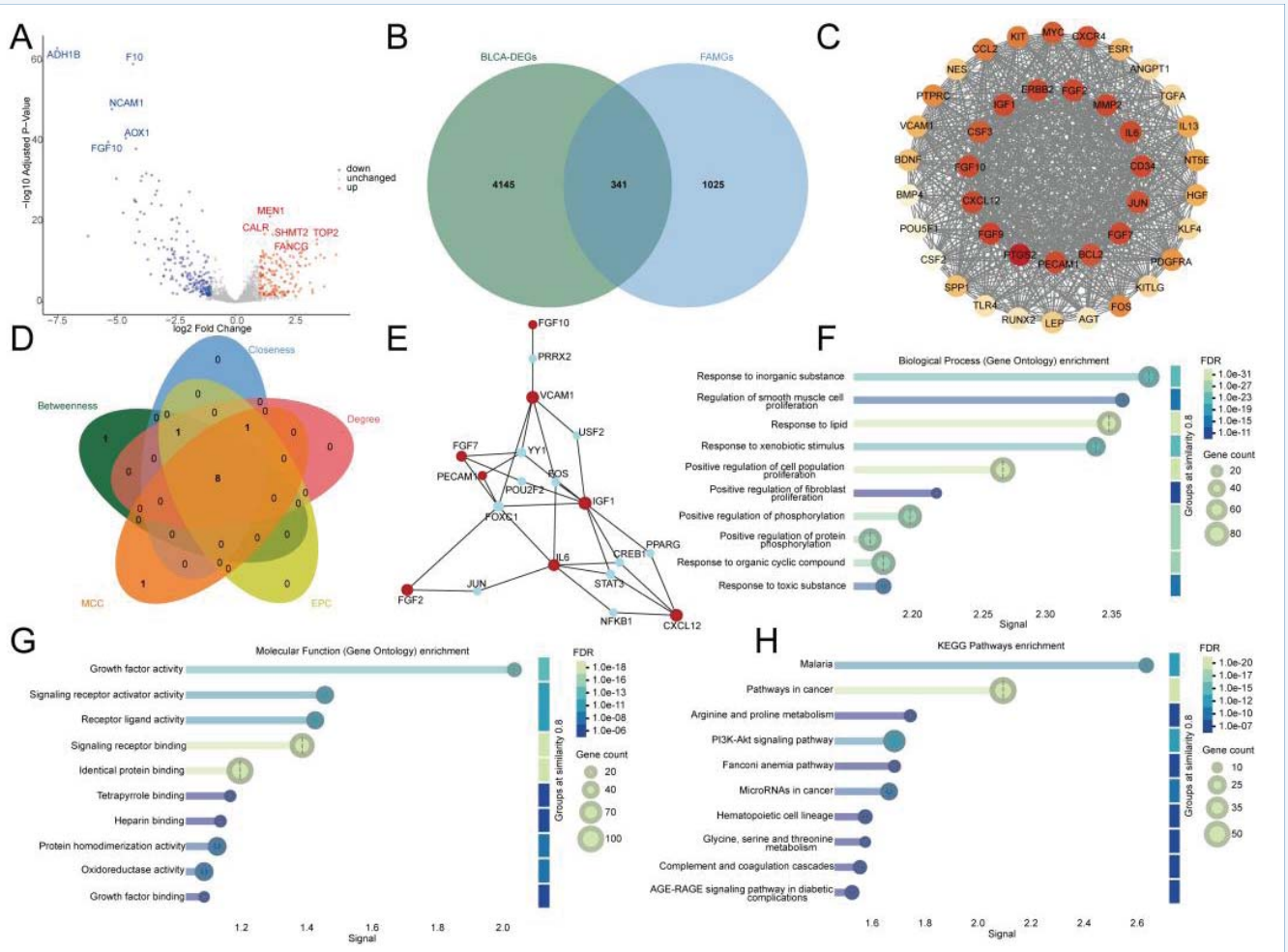

Figure 1 Identification of Differentially Expressed Folate Metabolism-Associated Genes (DEFAMGs). A) This volcano plot depicts the gene expression profiles that contrast bladder cancer tissues with adjacent non-cancerous samples. B) Venn diagram showing the intersection between BLCA-DEGs and FAMGs. C) STRING was used to generate a PPI network of overlapping genes, which was then visualized with Cytoscape. D) A Venn diagram in CytoHubba illustrates the hub genes identified using the algorithms MCC, EPC, Degree, Betweenness, and Closeness. E) PPI subnetwork comprising the eight hub genes. F) GO enrichment results for DEFAMGs within Biological Processes (BP). G) GO enrichment results for DEFAMGs within Molecular Functions (MF). H) KEGG pathway analysis of DEFAMGs.

As indicated by the Cumulative Distribution Function (CDF) curves and the delta area plot, the clustering reached optimal stability at k = 2, demonstrating strong uniformity within clusters and minimal variation between clusters (Figures 2A-C). The Kaplan-Meier analysis revealed that patients classified into cluster C2 exhibited a markedly poorer overall survival than those in cluster C1 (Figure 2D). According to the ESTIMATE algorithm, cluster C2 patients showed elevated stromal scores, immune scores, and ESTIMATE scores, accompanied by a markedly reduced tumor purity compared with those in cluster C1 (Figure 2E). According to the TIDE algorithm, patients in cluster C2 exhibited increased TIDE, T-cell impairment and rejection metrics (Figure 2F), pointing to an increased probability of immune system evasion within this subgroup. The ssGSEA analysis revealed distinct patterns of the infiltration of immune cells between the two groups, with cluster C2 exhibiting a notably higher overall level of immune infiltration than cluster C1 (Figures 2G,H).

Identification of the Hub Module and Genes Related to FAMG in Bulk RNA-Seq

To identify FAMG-associated gene modules linked to BLCA, a method for examining gene co-expression networks with Weighted

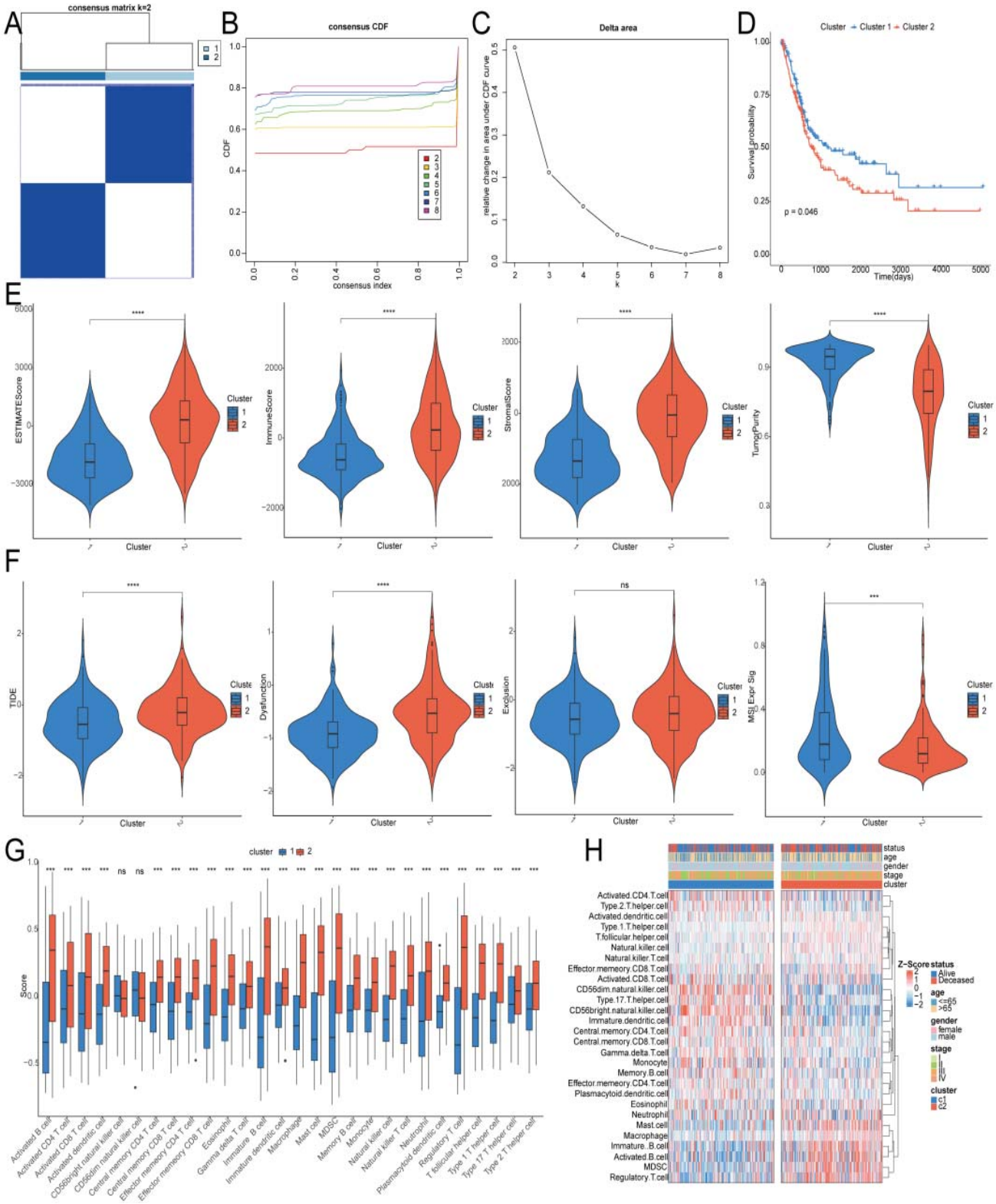

Figure 2 Identification of clusters related to FAMG-Related Genes. A) The consistency clustering analysis of 8 overlapping center DEFAMGs in the TCGA-BLCA cohort (k = 2). B) Color-coded CDF curves illustrating the consensus matrices for every k value. C) Optimal clustering stability at k = 2 is indicated by the relative changes in the area under the CDF curves for each k. D) Comparison of Kaplan–Meier overall survival curves for patients in clusters C1 and C2. E) Distribution of stromal, immune, and ESTIMATE scores, as well as tumor purity, across patient groups as determined by the ESTIMATE algorithm. F) TIDE analysis showing differences in TIDE, T-cell dysfunction, and T-cell rejection scores across clusters. G) The distribution of 28 immune-cell subsets across clusters C1 and C2. H) A heatmap displaying the relative abundance of 28 immune cell subsets assessed with the ssGSEA algorithm.

Associations (WGCNA) was conducted using the TCGA-BLCA cohort. A soft-thresholding power of 3 was selected to achieve a network without a characteristic scale topology (Figure 3A). Eleven co-expression modules were obtained, each visualized in a distinct color to represent different gene clusters. The turquoise module showed the strongest correlation with BLCA clusters among all identified modules, with a correlation coefficient of 0.73 and a *p*-value less than 0.001 (Figures 3B,C). To further explore the biological significance, we intersected the turquoise module genes with FAMG-DEGs, yielding 74 overlapping key genes (Figure 3D). Using the expression data of the 74 overlapping genes, univariate Cox regression identified 27 genes as prognostic markers that are significantly related to overall survival (Figure 3E).

Development and validation of the FAMPI

We applied four machine learning algorithms-LASSO Cox regression, XGBoost, decision tree, and random forest-to further

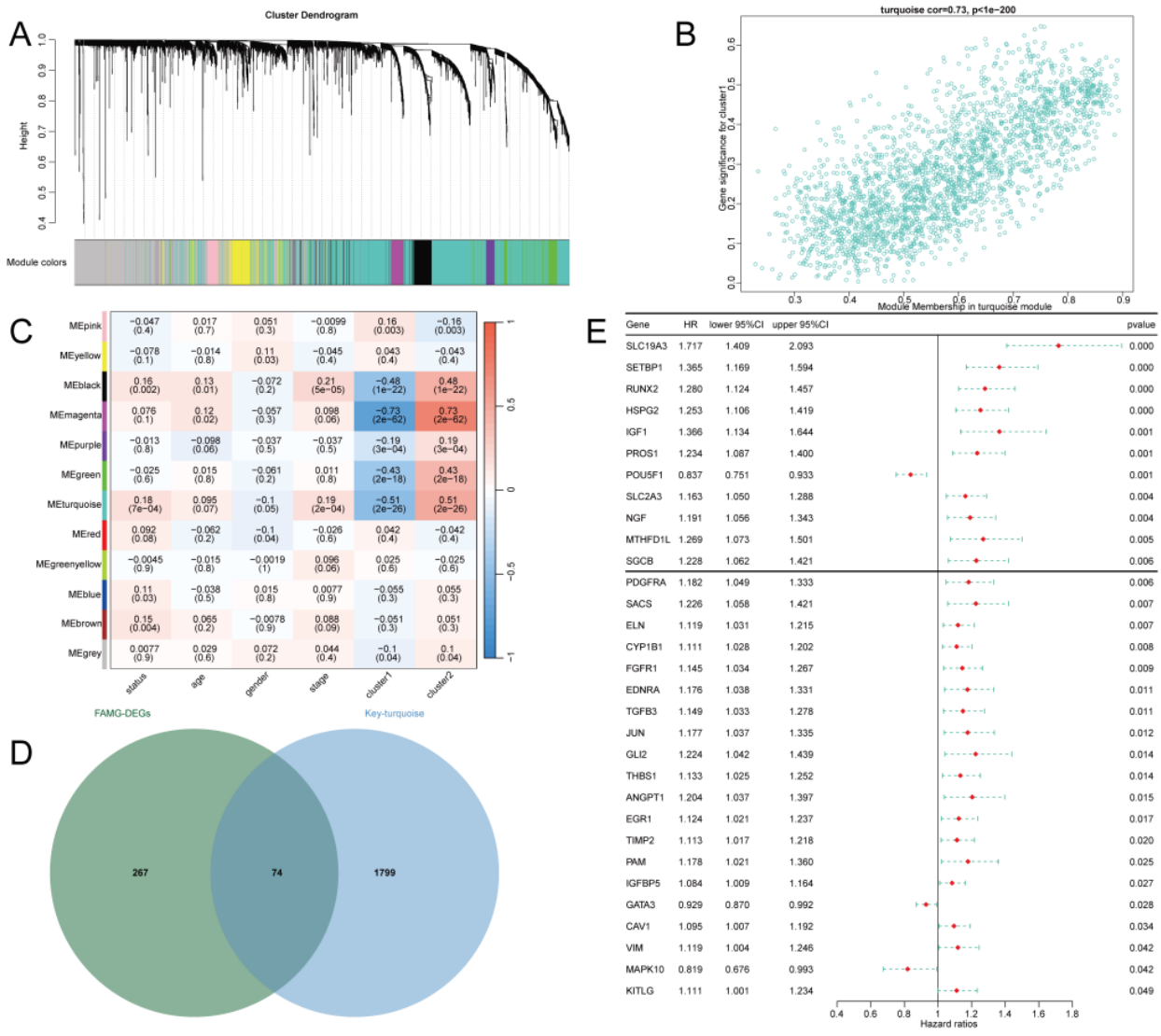

Figure 3 Identification of the Hub Module and Genes Related to FAMG in Bulk RNA-Seq. A) Hierarchical clustering dendrogram of genes constructed by WGCNA, where different colors represent distinct modules. B) Relationship between Gene Significance (GS) and Module Membership (MM) for genes in the turquoise module, visualized in a scatter plot. C) Correlation heatmap between module eigengenes and clinical traits in BLCA patients. D) The intersection between the turquoise module genes and FAMG-DEGs, revealing a core set of 74 overlapping genes. E) Univariate Cox analysis in the TCGA-BLCA cohort discovered 27 prognostic DEFAMGs.

screen key prognostic genes from the 27 prognostic DEFAMGs. Consequently, six representative genes (SLC19A3, MTHFD1L, CAV1, POU5F1, SETBP1, and HSPG2) were identified (Figure 4A). Based on the expression patterns and regression coefficients of these six genes, we established a folate metabolism–associated prognostic index (FAMPI) as follows: FAMPI = (SLC19A3 × 0.43465778) + (MTHFD1L × 0.22104801) + (CAV1 × −0.11899016) + (POU5F1 × − 0.12543035) + (SETBP1 × 0.28322608) + (HSPG2 × 0.05826853). BLCA patients were divided into high and low FAMPI groups based on the median FAMPI score. Kaplan-Meier survival analysis indicated that patients with high FAMPI scores had much worse overall survival than those with low scores in both the TCGA training cohort ($p < 0.001$, Figure 4B) and the independent GSE13507 validation cohort ($p = 0.0035$, Figure 4C). A prognostic nomogram integrating FAMPI with age, gender, and clinical stage was established to estimate the overall survival probabilities for 1, 3, and 5 years (Figure 4D). Cox regression analyses, encompassing both univariate and multivariate models, revealed that FAMPI functions as an independent prognostic biomarker in BLCA (Figures 4E,F). The calibration plots indicated a close agreement between the predicted and observed survival outcomes (Figure 4G). Moreover, the decision curve analysis illustrated that the nomogram provided superior clinical

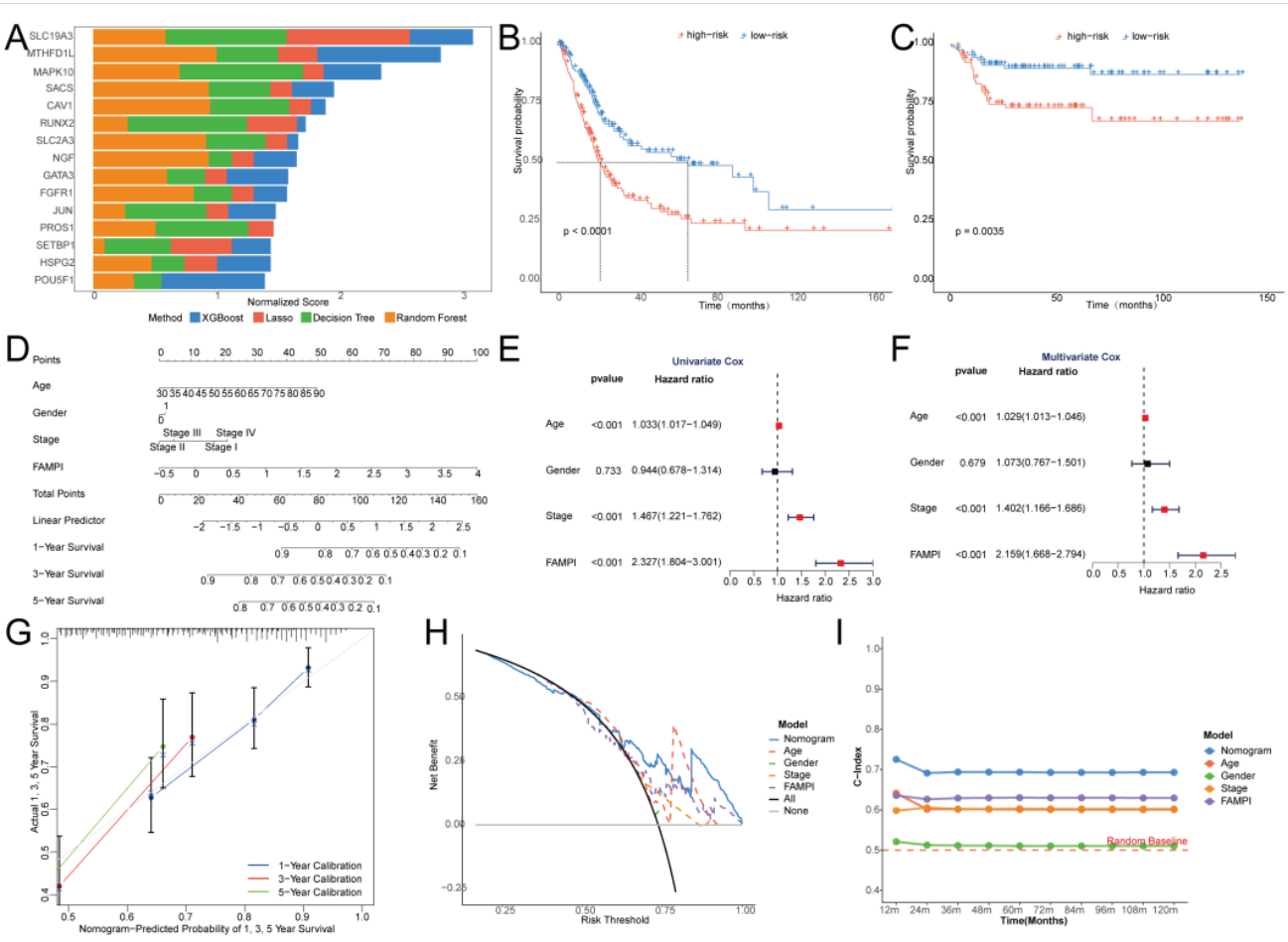

Figure 4 Development and validation of the FAMPI. A) The top 15 genes identified using four machine learning algorithms. B) Kaplan-Meier survival plots for the training cohort of TCGA-BLCA. C) Kaplan–Meier survival plots for the independent validation cohort GSE13507. D) Construction of the nomogram integrating FAMPI with clinical characteristics. E) Univariate Cox regression analysis. F) Multivariate Cox regression analysis. G) Calibration plots of the nomogram predicting 1-, 3-, and 5-year Overall Survival (OS) in the TCGA-BLCA cohort. H) Decision Curve Analysis (DCA) illustrating the net clinical benefit derived from the application of the nomogram compared with conventional clinical variables. I) Comparative analysis of the C-index between the nomogram and traditional clinical parameters.

utility over traditional clinicopathological parameters (Figure 4H). In addition, the C-index confirmed its superior discriminative performance compared with other evaluated models (Figure 4I). Collectively, these results highlight that the FAMPI-based nomogram represents a reliable and individualized tool for prognostic assessment in BLCA patients.

The accuracy of this nomogram's predictions was confirmed through ROC analysis. In the TCGA-BLCA training group, the AUC values for overall survival at 1, 3, and 5 years were 0.73, 0.73, and 0.75, respectively (Figure 5A). Consistent findings were obtained in the external validation cohort GSE13507, where the AUCs reached 0.95, 0.95, and 0.97 (Figure 5B). As shown in figures 5C,D the FAMPI distribution, patient survival outcomes, and expression patterns of the six FAMPI genes varied markedly between the two FAMPI subgroups. Among the six genes, MTHFD1L, SETBP1, SLC19A3, and HSPG2 showed positive coefficients and acted as risk factors, whereas CAV1 and POU5F1 exhibited negative coefficients and were associated with a protective effect. In the high-FAMPI group, all six genes showed significantly altered expression levels compared with the low-FAMPI group, suggesting that this integrated signature captures a dynamic balance between risk-enhancing and protective

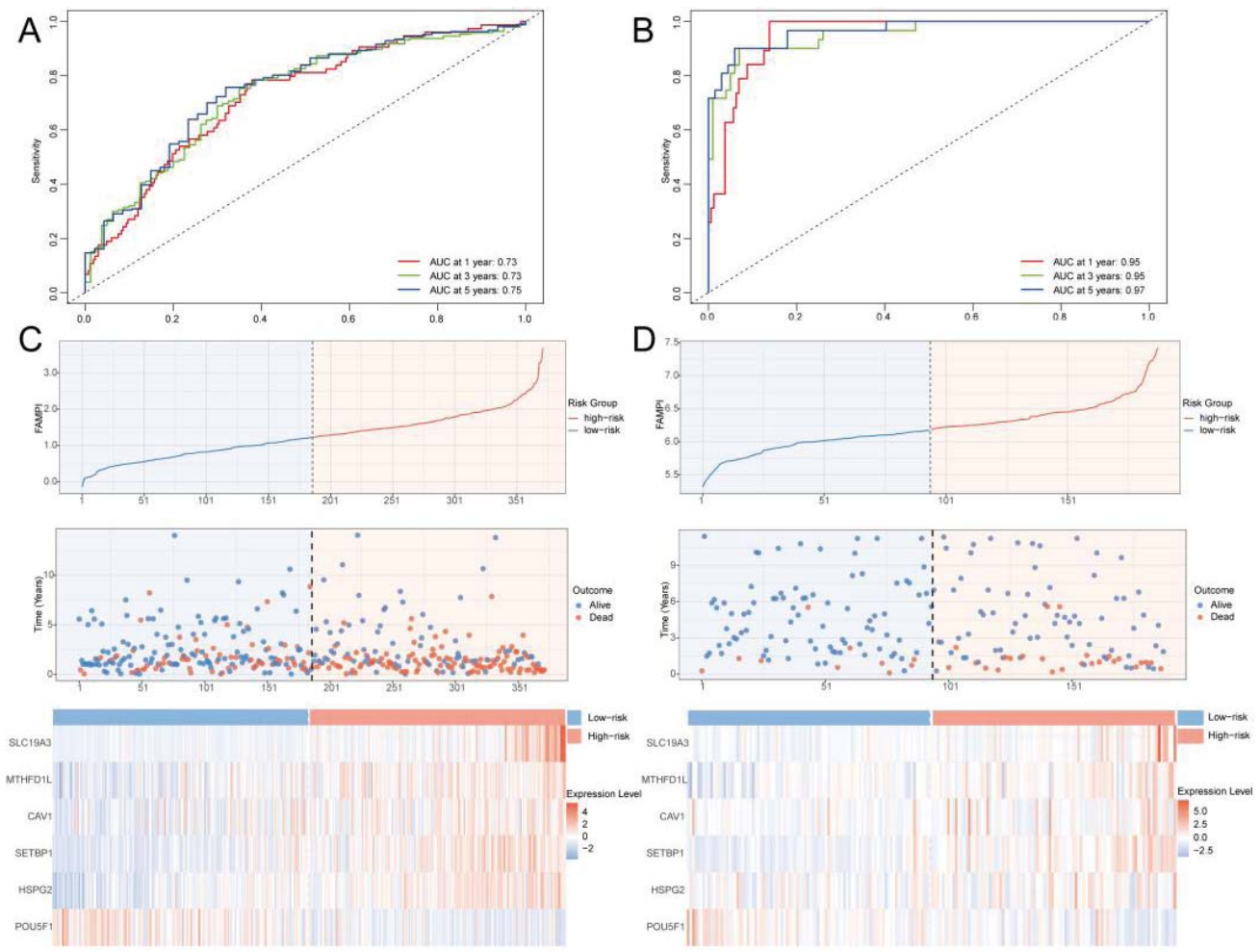

Figure 5 The prognostic value of FAMPI. A) ROC curves illustrating the predictive accuracy of FAMPI for 1-, 3-, and 5-year Overall Survival (OS) in the TCGA-BLCA cohort. B) ROC curves depicting the predictive performance of FAMPI for 1-, 3-, and 5-year OS in the independent validation cohort GSE13507. C) Distribution of FAMPI scores, patient survival outcomes, and expression patterns of the six FAMPI genes (SLC19A3, MTHFD1L, CAV1, SETBP1, HSPG2, and POU5F1) in the TCGA-BLCA cohort. D) Distribution of FAMPI scores, patient survival outcomes, and expression patterns of the six FAMPI genes in the GSE13507 cohort.

genes. This trend was further substantiated in the independent validation dataset GSE13507. Collectively, these observations demonstrate that FAMPI possesses strong predictive efficacy for both short- and long-term survival outcomes in BLCA patients.

Clinical Relevance Assessment of the FAMPI

Prognostic significance of the FAMPI signature in subgroups with distinct clinical profiles. Based on survival status, molecular cluster classification, risk score, age, gender, and pathological stage, patients were categorized into different subgroups for comparison. The findings indicated that FAMPI expression levels were markedly increased in non-survivors, in cluster 2, in the high-risk group, in individuals older than 60 years, and in those at advanced pathological stages III–IV (Figures 6A-F). Moreover, Kaplan–Meier survival analysis consistently demonstrated that patients with higher FAMPI scores had considerably poorer overall survival than those with lower scores across different subgroups, including those aged 60 years or younger, those older than 60 years, females, males, stages I–II, and stages III–IV (Figures 6G-L). Collectively, these results underscore that FAMPI functions as a robust and stable prognostic biomarker for BLCA patients across diverse clinical contexts.

Gene set enrichment and tumor immune microenvironment analysis

Gene Set Enrichment Analysis (GSEA) was conducted using Gene Ontology (GO) terms to explore the biological mechanisms that differentiate the two risk groups. The analysis revealed that high-risk patients were significantly enriched in pathways associated with cell growth, cellular response to growth factor stimulation, and T-cell activation, whereas low-risk patients showed enrichment in processes such as amine transport, cellular response to nitrogen compounds, and homeostatic regulation (Figures 7A,B). Considering the close association between the risk score and the tumor immune

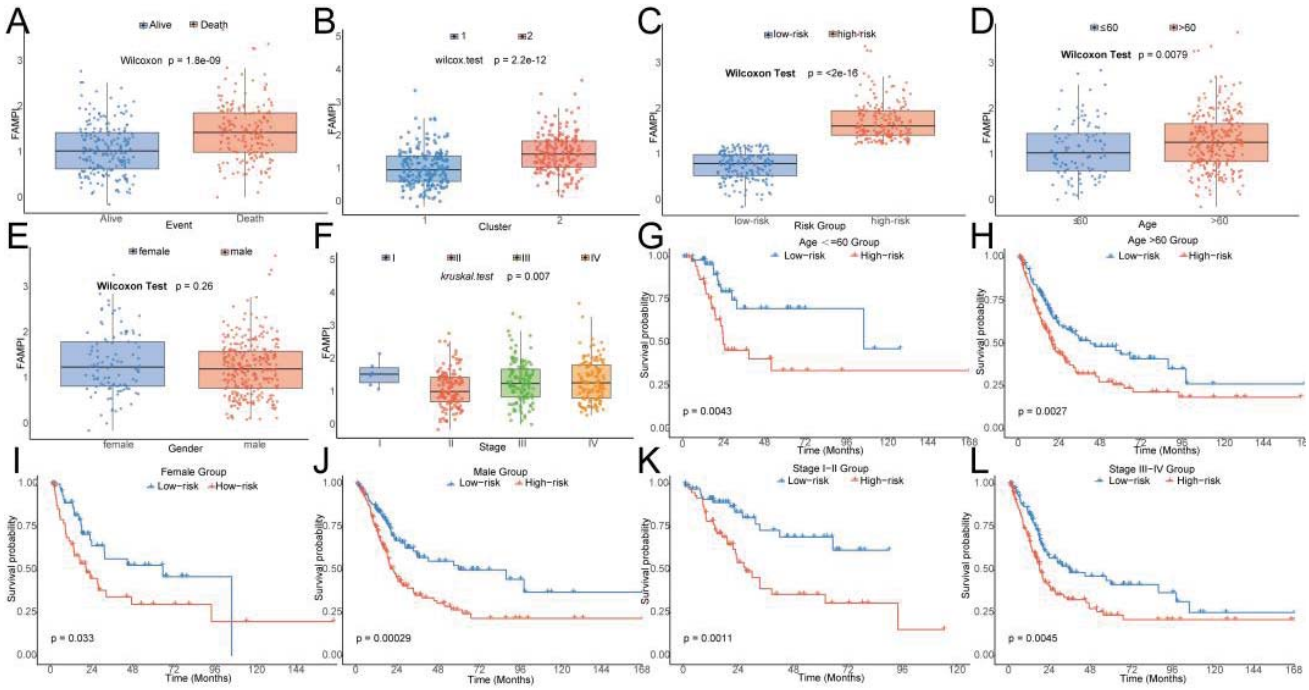

Figure 6 Clinical relevance of the FAMPI signature in bladder cancer patients. A-F) Boxplots showing FAMPI score distributions according to clinical features: survival status (alive vs., death), molecular cluster (C1 vs., C2), risk group (low- vs., high-risk), age (≤ 60 vs., > 60 years), gender (female vs., male), and pathological stage (I-IV). G-L) Kaplan–Meier overall survival curves comparing high- and low-risk groups within subgroups stratified by age (≤ 60 vs., > 60 years), gender (female vs. male), and pathological stage (I-II vs., III-IV).

microenvironment, subsequent immune evaluation was conducted to assess infiltration using multiple computational algorithms. The high-risk cohort demonstrated significantly elevated immune, stromal, and ESTIMATE scores, while exhibiting reduced tumor purity in comparison to those in the low-risk cohort (Figure 7C). Further immune profiling revealed that most immune cell types differed markedly between the two groups, with the high-risk cohort showing elevated infiltration of a wide range of immune cells (Figure 7D). Moreover, correlation analysis revealed significant covariation between the six FAMPI genes and levels of immune cell infiltration, highlighting a dynamic interplay between the expression of these genes and the composition of the tumor immune landscape (Figure 7E). These findings imply that the convergence of aberrant biological processes and a distinctive immune microenvironment collectively drives the adverse clinical outcomes in high-risk patients.

Genomic variation landscape

To characterize mutational distinctions between FAMPI categories, we compared the somatic mutation landscapes of the high- and low-FAMPI cohorts. TP53, TTN, KMT2D, ARID1A, and MUC16 were the top recurrently altered genes in both cohorts, with a slightly elevated mutation frequency in the high-FAMPI cohort. (Figures 8A,B). Analysis of co-

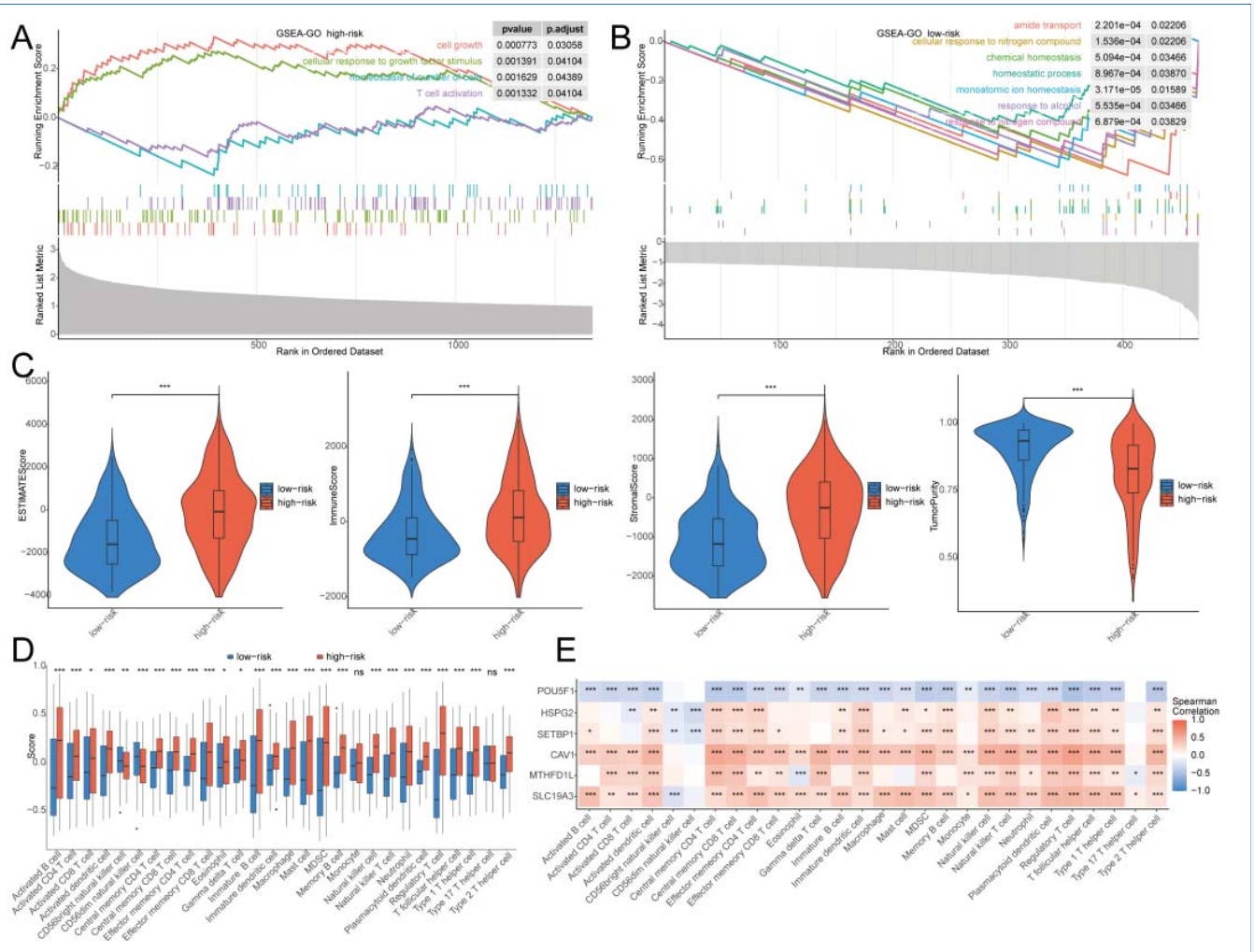

Figure 7 Gene set enrichment analysis and tumor immune microenvironment characteristics between FAMPI-defined risk groups. A-B) GSEA based on GO terms comparing high- and low-risk groups. C) Violin plots showing differences in ESTIMATE scores, immune scores, stromal scores, and tumor purity between the two risk groups. D) Boxplots of the 28 tumor-infiltrating immune cell subsets in high- and low-risk groups. E) Heatmap showing correlations between the six FAMPI genes and tumor-infiltrating immune cell populations.

occurrence and mutual exclusivity revealed that most mutations exhibited cooperative patterns, while only a few showed exclusionary effects. Specifically, KMT2A–TP53 and ARID1A–RB1 mutations tended to co-occur in the high-FAMPI group, whereas mutations in MUC16 and TTN showed a moderate tendency for co-occurrence in the low-FAMPI cohort (Figures 8C,D). Additionally, mutation profiling of the six FAMPI genes (SLC19A3, MTHFD1L, CAV1, SETBP1, HSPG2, and POU5F1) demonstrated that most alterations were dominated by missense mutations, with occasional nonsense and splice site variants (Figure 8E).

Prediction of immunotherapy efficacy and immune function

Immune Checkpoint Inhibitors (ICIs) have demonstrated remarkable therapeutic efficacy across multiple malignancies. In this analysis, patients classified in the high-FAMPI cohort displayed significantly elevated TIDE values and enhanced T-cell exclusion indices. In contrast, T-cell dysfunction and MSI levels showed no notable differences between the subgroups, implying an increased potential for immune escape in individuals with high-FAMPI BLCA (Figure 9A). Similarly, IPS-based assessment indicated an absence of notable differences across the majority of immune subtypes between the high- and low-FAMPI cohorts (Figure 9B). Moreover, ssGSEA assessment revealed that several immune-associated pathways—including MHC class I antigen presentation, APC co-stimulatory signaling, immune checkpoint regulation, and T-cell activation—were more prominently active within the high-FAMPI cohort (Figure 9C). Correlation analysis further demonstrated that the six representative prognostic genes were significantly and positively associated with the majority of immune checkpoint molecules (Figure 9D). Additionally, the majority of immune checkpoint molecules—including CD274 (PD-L1), PDCD1 (PD-1), CTLA4, LAG3,

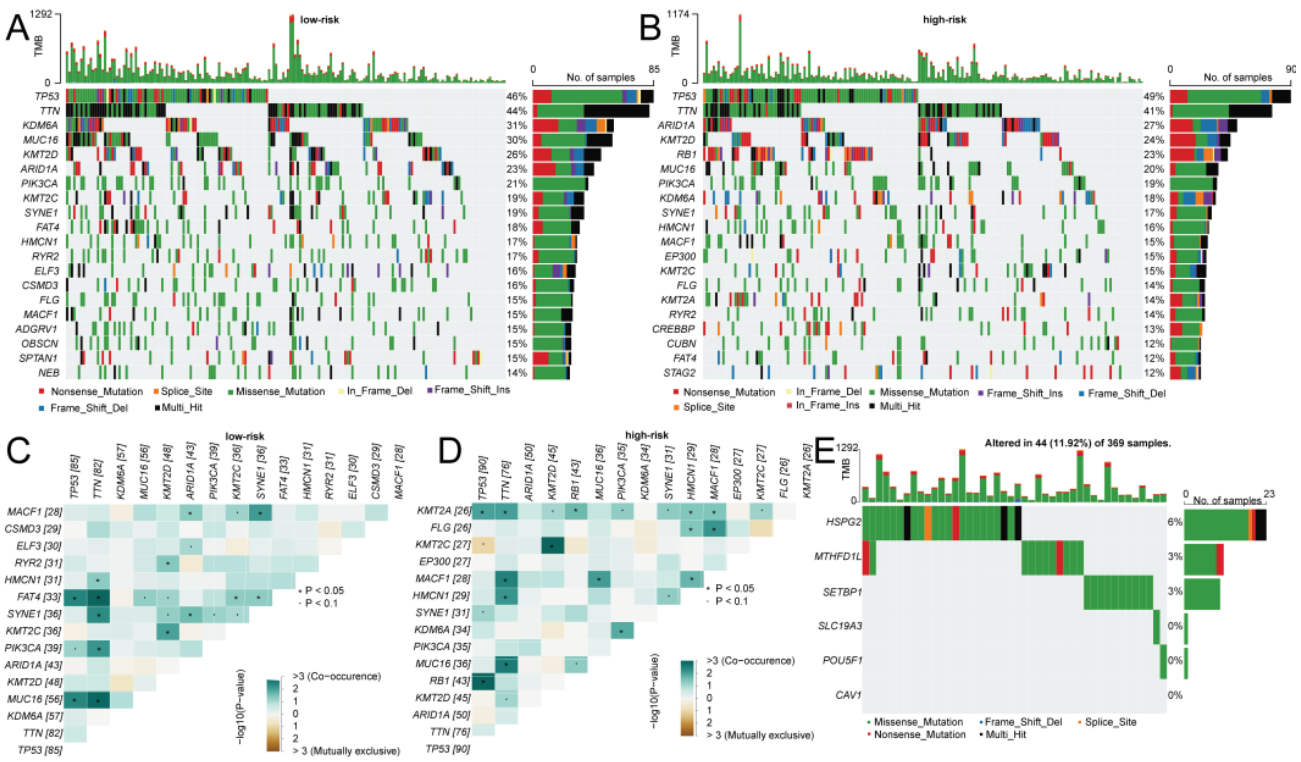

Figure 8 Genomic variation landscape between the two FAMPI groups. Waterfall plots illustrating somatic mutation profiles in the low-FAMPI group (A) and the high-FAMPI group (B). Heatmaps showing mutually exclusive and co-occurring gene mutations in the low-FAMPI group (C) and the high-FAMPI group (D). (E) Mutation frequencies of six representative FAMPI-related genes.

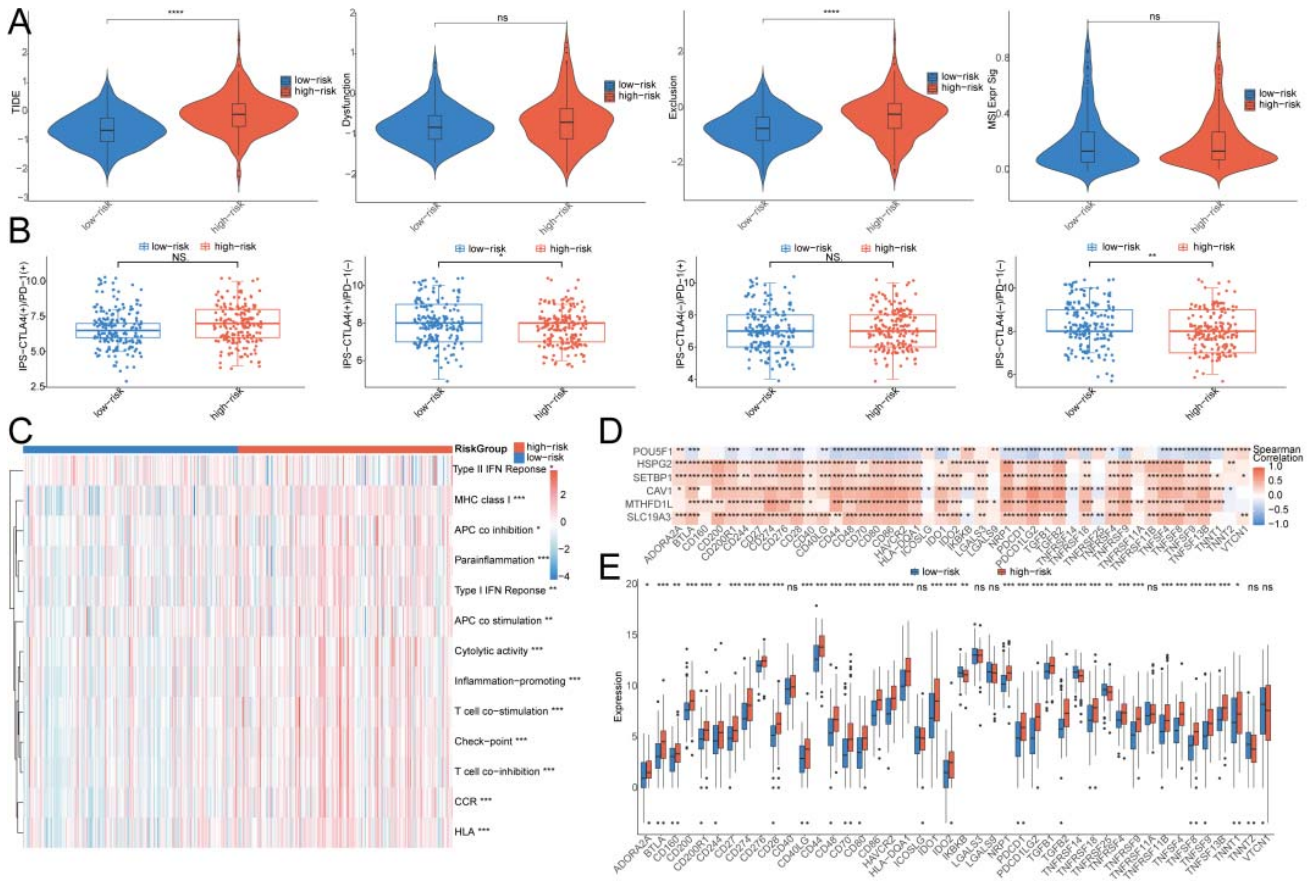

Figure 9 Prediction of immunotherapy efficacy and immune function. A) Comparisons of TIDE, T-cell dysfunction, T-cell exclusion, and MSI scores between the high- and low-FAMPI groups. B) IPS scores across different subtypes in the two FAMPI groups. C) Activities of immune-associated pathways showing significant differences between the high- and low-FAMPI cohorts based on ssGSEA. D) Correlations between six representative prognostic genes and immune checkpoint regulators. E) Immune checkpoint–related gene expression patterns across high- and low-FAMPI cohorts.

TIGIT, HAVCR2 (TIM-3), CD276 (B7-H3), and ICOS—were markedly elevated in the high-FAMPI cohort, while a few, for instance TMIGD2, showed no notable differences (Figure 9E).

Drug Sensitivity Analysis

To further explore the clinical relevance of the six FAMPI-associated prognostic genes in facilitating personalized therapy for BLCA, we assessed the responsiveness of frequently applied chemotherapeutic agents within the high- and low-FAMPI cohorts. According to the IC50 results, individuals in the high-FAMPI group displayed markedly reduced IC50 values for several standard chemotherapeutic agents, suggesting that the FAMPI score may serve as a predictive marker of chemotherapy responsiveness in BLCA patients (Figure 10A). To further investigate the association between the expression profiles of the six significant prognostic genes and drug sensitivity, we employed the CellMiner platform for analysis. Correlation plots of the five drugs showing the most pronounced correlations indicated that positive correlations were linked to elevated gene expression and enhanced drug sensitivity, while negative correlations implied greater drug resistance (Figure 10B). In addition, data obtained from the GDSC database further supported that expression levels of certain characteristic genes were positively correlated with multiple drug sensitivities, whereas others showed negative correlations, reinforcing their potential roles as biomarkers for chemotherapy response in BLCA (Figure 10C).

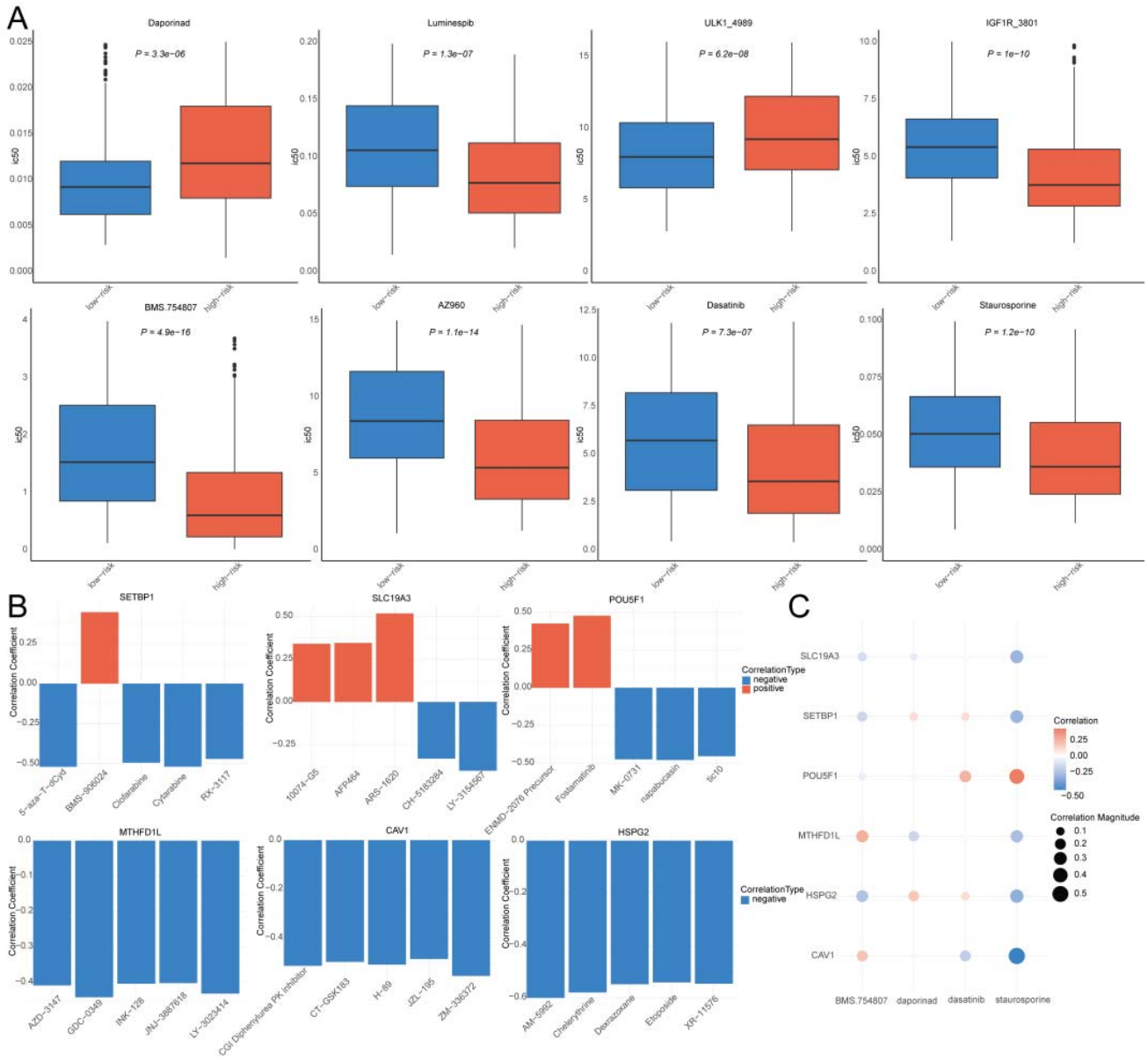

Figure 10 Drug sensitivity analysis between the two FAMPI groups. A) Illustrates the IC50 comparison of frequently applied anticancer drugs in the high- and low-FAMPI subgroups. B) Presents the association between the expression profiles of pivotal characteristic genes in the FAMPI model and drug sensitivity, with bar plots representing the top five drugs most strongly associated with each gene. C) Presents a correlation heatmap linking the six FAMPI genes of the FAMPI model to drug sensitivity data obtained from the GDSC database.

DISCUSSION

Folate metabolism is essential for DNA synthesis and methylation, thereby affecting tumor cell proliferation and therapeutic response [51,52]. Clinically, DHFR and TYMS have become classical antimetabolite targets, highlighting the translational potential of Folate Metabolism-Associated Genes (FAMGs) [53]. In this study, we systematically profiled Folate Metabolism–Associated Genes (FAMGs) in bladder cancer (BLCA). Differential expression analysis identified 341 DEFAMGs (the intersection of BLCA-DEGs and FAMGs, Figure 1B), among which 194 were upregulated and 147 were downregulated. Consensus clustering divided patients into two molecular subtypes, among which Cluster 1 exhibited the highest tumor purity and the lowest immune cell infiltration, stromal content, and ESTIMATE index, together with the highest T-cell exclusion

and impairment scores, and reduced MSI. These findings indicate an immunologically "cold" but immune-evasive phenotype. This is consistent with prior evidence that MTHFD2 can upregulate CD274 expression through IFN-γ signaling [54], supporting the concept that metabolic reprogramming can shape the immune microenvironment in BLCA.

To investigate the molecular factors contributing to these phenotypic differences, we applied a Weighted Gene Co-Expression Network Approach (WGCNA). The turquoise module, which showed the strongest negative correlation with Cluster 1, was identified as the most clinically relevant. From this module, we refined 74 core genes overlapping with DEFAMGs. Subsequent univariate Cox regression and machine learning analyses identified six key genes (SLC19A3, MTHFD1L, SETBP1, CAV1, POU5F1 and HSPG2) to construct the FAMPI risk signature. These genes encompassed both oncogenic drivers and potential protective factors, highlighting a balanced interplay between metabolism, immunity, and prognosis. For instance, MTHFD1L, a mitochondrial one-carbon enzyme, has been recognized as a poor prognostic indicator in colorectal cancer [55,56] and hepatocellular carcinoma [57], and we verified its contribution as a risk factor in BLCA [58]. Similarly, SETBP1 is a recognized oncogene promoting proliferation [59], significantly overexpressed in the high-risk group. Conversely, CAV1, known for its context-dependent dual roles [60], showed a protective effect in our model. Intriguingly, POU5F1 (Oct4), though generally linked to stemness and aggressiveness, acted as a favorable factor here, in line with tumor-type–specific heterogeneity [61]. Finally, SLC19A3, though not a classical folate transporter, has been implicated in drug sensitivity [62], suggesting a role in modulating therapeutic response in BLCA. In addition, HSPG2, encoding perlecan—a key heparan sulfate proteoglycan within the extracellular matrix—has been associated with tumor invasion and angiogenic processes [63], and was recognized as a risk factor in our model.

From a clinical perspective, FAMPI demonstrated superior prognostic predictive performance compared with traditional clinical features across both TCGA and GSE13507 cohorts. Notably, high-FAMPI patients exhibited an immunosuppressive Tumor Microenvironment (TME) with unregulated immune Checkpoints (CD274, CTLA4) and higher Immunophenoscores (IPS). Despite low-FAMPI patients having higher TIDE scores, suggesting immune dysfunction, this does not conflict with our conclusion. TIDE primarily reflects T-cell dysfunction or exclusion, while IPS and immune checkpoint expression represent an "inflamed but suppressed" state [64]. Clinically, such a state is often associated with greater sensitivity to Immune Checkpoint Inhibitors (ICIs) [65]. Therefore, FAMPI not only stratifies prognosis but also predicts immunotherapy benefit, offering value for patient selection. In addition, differential drug sensitivity analyses suggested that FAMPI may guide chemotherapy choice, further supporting its clinical utility.

Mechanistically, pathway enrichment analysis suggested that the high-risk group exhibited notable enrichment of TGF-β, BMP, and PI3K–Akt signaling cascades, which participate in immune suppression, fibroblast activation, and tumor growth. The enrichment of TGF-β signaling in particular provides a plausible explanation for the immunosuppressive phenotype [66]. Conversely, the low-risk group demonstrated marked enrichment in core metabolic pathways, notably ribosome biogenesis and oxidative phosphorylation. Furthermore, distinct mutational patterns between groups—including alterations in Chromatin Remodeling–Related Genes (e.g., KMT2A, CREBBP) that were more frequently observed within the high-risk cohort, whereas alterations such as MACF1 mutations were identified in the low-risk cohort—suggest that folate metabolic reprogramming may collaborate with genomic instability to drive divergent BLCA trajectories [67]. Collectively, these findings elucidate the mechanistic basis

underlying the observed immunological and clinical heterogeneity.

Importantly, compared with previously reported immune or metabolism-based models [68,69], FAMPI represents the first folate metabolism–focused prognostic index in BLCA. It robustly integrates metabolic reprogramming, immune microenvironment characteristics, somatic mutation landscapes, and therapeutic response prediction. Thus, FAMPI functions not merely as a prognostic marker but also as a promising instrument for decision support in guiding precision treatment strategies, thereby promoting advances in personalized oncology.

In addition, the predictive performance of FAMPI for overall survival was strongly validated. Within the TCGA dataset, the AUCs for 1-, 3-, and 5-year overall survival reached 0.77, 0.75, and 0.73, respectively. In the independent GSE13507 dataset, the corresponding AUCs reached 0.95, 0.93, and 0.91, confirming the strong robustness and reproducibility of the model. Furthermore, the drug sensitivity analysis uncovered significant variations in therapeutic responses to commonly used chemotherapeutic agents across the different risk groups. Notably, high expression of SETBP1 was consistently associated with poorer outcomes and chemotherapy resistance. Interestingly, For POU5F1, although it acted as a favorable prognostic factor within the FAMPI model, its high expression also correlated with distinct drug response patterns. This suggests that the role of POU5F1 may be context-dependent, reflecting tumor type–specific heterogeneity and potentially divergent mechanisms in prognosis versus therapy sensitivity. These findings not only reinforce the clinical value of FAMPI in prognostic prediction but also highlight its potential as a supportive tool for guiding individualized chemotherapy selection.

However, the conclusions of this study are mainly based on data analysis from public databases, and several limitations should be noted: (1) the sample size collected in this study was limited, and additional clinical samples from larger cohorts with expanded study populations are required to enhance the credibility of the findings. (2) The transcriptomic and clinical datasets of BLCA patients in this study were primarily obtained from the TCGA and GEO public databases, which may introduce differences between datasets; thus, further efforts are needed to address potential data bias and improve the reliability of the analysis. (3) The analytical results of this study still require validation in broader patient populations, and additional molecular experiments are necessary to further elucidate the molecular mechanisms of FAMPI in BLCA patients. (4) The drug sensitivity profiles associated with the FAMPI signature were inferred computationally using in silico prediction tools based on transcriptomic data; these findings must be corroborated by functional assays in BLCA cell lines and other experimental models before they can be translated into clinical decision-making.

CONCLUSION

In summary, the FAMPI model established and validated herein can serve not merely as a potential molecular classifier for individuals with bladder cancer but also as a dependable indicator of therapeutic responses to chemotherapy and immunotherapy. Moreover, it is expected to provide valuable insights for personalized clinical management and treatment decision-making in BLCA.

FUNDING

This work was supported by grants from the Guangxi Natural Science Foundation (2024GXNSFAA010096), Guangdong Basic and Applied Basic Research Foundation (2025A1515012661), Guangdong Province Medical Science and Technology Research Fundation (B2025207), Guangxi Medical and Health Appropriate Technology Development and Promotion Project (S2024075), Innovation Training Program for College Students (S202510601079, X202510601230, X202510601258), Guangxi Key Laboratory of

Tumor Immunology and Microenvironmental Regulation (2023KF006, 3030302213).

AVAILABILITY OF DATA AND MATERIALS

The data generated in the present study may be requested from the corresponding author.

AUTHORS' CONTRIBUTION

YM and BZ designed the study, performed data curation and formal analysis, and wrote the original draft. YL and FY participated in data analysis and validation. CL assisted with methodology and investigation. JH contributed to data interpretation and manuscript review. XH supervised the study, acquired funding, conceptualized the research, and reviewed and edited the manuscript. All authors read and approved the final manuscript.

COMPETING INTERESTS

The authors declare that they have no competing interests.

REFERENCES

1. Siegel RL, Miller KD, Wagle NS, Jemal A. Cancer statistics. CA Cancer J Clin. 2023;73(1):17-48. doi: 10.3322/caac.21763.
2. Tran L, Xiao JF, Agarwal N, Duex JE, Theodorescu D. Advances in bladder cancer biology and therapy. Nat Rev Cancer. 2021 Feb;21(2):104-121. doi: 10.1038/s41568-020-00313-1. Epub 2020 Dec 2. PMID: 33268841; PMCID: PMC10112195.
3. Wucherpfennig S, Rose M, Maurer A, Cassataro MA, Seillier L, Morsch R, Hammad E, Baldia PH, Ecke TH, Vögeli TA, Knüchel R, Gaisa NT. Evaluation of Therapeutic Targets in Histological Subtypes of Bladder Cancer. Int J Mol Sci. 2021 Oct 26;22(21):11547. doi: 10.3390/ijms222111547. PMID: 34768978; PMCID: PMC8583926.
4. Saginala K, Barsouk A, Aluru JS, Rawla P, Padala SA, Barsouk A. Epidemiology of bladder cancer. Med Sci (Basel). 2020;8(1):15. doi: 10.3390/medsci8010015.
5. Babjuk M, Burger M, Capoun O, Cohen D, Compérat EM, Dominguez Escrig JL, Gontero P, Liedberg F, Masson-Lecomte A, Mostafid AH, Palou J, van Rhijn BWG, Rouprêt M, Shariat SF, Seisen T, Soukup V, Sylvester RJ. European Association of Urology Guidelines on Non-muscle-invasive Bladder Cancer (Ta, T1, and Carcinoma in Situ). Eur Urol. 2022 Jan;81(1):75-94. doi: 10.1016/j.eururo.2021.08.010. Epub 2021 Sep 10. PMID: 34511303.
6. Bellmunt J, de Wit R, Vaughn DJ, Fradet Y, Lee JL, Fong L, Vogelzang NJ, Climent MA, Petrylak DP, Choueiri TK, Necchi A, Gerritsen W, Gurney H, Quinn DI, Culine S, Sternberg CN, Mai Y, Poehlein CH, Perini RF, Bajorin DF; KEYNOTE-045 Investigators. Pembrolizumab as Second-Line Therapy for Advanced Urothelial Carcinoma. N Engl J Med. 2017 Mar 16;376(11):1015-1026. doi: 10.1056/NEJMoa1613683. Epub 2017 Feb 17. PMID: 28212060; PMCID: PMC5635424.
7. Yajima S, Masuda H. Immune checkpoint inhibitors and antibody-drug conjugates in urothelial carcinoma: Current landscape and future directions. Cancers (Basel). 2025;17(9):1594. doi: 10.3390/cancers17091594.
8. Dagogo-Jack I, Shaw AT. Tumour heterogeneity and resistance to cancer therapies. Nat Rev Clin Oncol. 2018 Feb;15(2):81-94. doi: 10.1038/nrclinonc.2017.166. Epub 2017 Nov 8. PMID: 29115304.
9. Sobral AF, Cunha A, Silva V, Gil-Martins E, Silva R, Barbosa DJ. Unveiling the Therapeutic Potential of Folate-Dependent One-Carbon Metabolism in Cancer and Neurodegeneration. Int J Mol Sci. 2024 Aug 28;25(17):9339. doi: 10.3390/ijms25179339. PMID: 39273288; PMCID: PMC11395277.
10. Newman AC, Maddocks ODK. One-carbon metabolism in cancer. Br J Cancer. 2017 Jun 6;116(12):1499-1504. doi: 10.1038/bjc.2017.118. Epub 2017 May 4. PMID: 28472819; PMCID: PMC5518849.
11. Ren X, Wang X, Zheng G, Wang S, Wang Q, Yuan M, Xu T, Xu J, Huang P, Ge M. Targeting one-carbon metabolism for cancer immunotherapy. Clin Transl Med. 2024 Jan;14(1):e1521. doi: 10.1002/ctm2.1521. PMID: 38279895; PMCID: PMC10819114.
12. Zhang T, Liu J, Wang M, Liu X, Qu J, Zhang H. Prognosis stratification and response to treatment in breast cancer based on one-carbon metabolism-related signature. Front Oncol. 2024 Jan 4;13:1288909. doi: 10.3389/fonc.2023.1288909. PMID: 38239641; PMCID: PMC10794736.
13. Zhu Y, Zhou T, Zheng Y, Yao Y, Lin M, Zeng C, Yan Y, Zhou Y, Li DD, Zhang J. Folate metabolism-associated CYP26A1 is a clinico-immune target in colorectal cancer. Genes Immun. 2025 Aug;26(4):376-393. doi: 10.1038/

s41435-025-00342-6. Epub 2025 Jul 3. PMID: 40604317; PMCID: PMC12353796.

14. Cancer Genome Atlas Research Network; Weinstein JN, Collisson EA, Mills GB, Shaw KR, Ozenberger BA, Ellrott K, Shmulevich I, Sander C, Stuart JM. The Cancer Genome Atlas Pan-Cancer analysis project. Nat Genet. 2013 Oct;45(10):1113-20. doi: 10.1038/ng.2764. PMID: 24071849; PMCID: PMC3919969.

15. Law CW, Chen Y, Shi W, Smyth GK. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biol. 2014 Feb 3;15(2):R29. doi: 10.1186/gb-2014-15-2-r29. PMID: 24485249; PMCID: PMC4053721.

16. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;43(7):e47. doi: 10.1093/nar/gkv007.

17. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P, Jensen LJ, Mering CV. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. 2019 Jan 8;47(D1):D607-D613. doi: 10.1093/nar/gky1131. PMID: 30476243; PMCID: PMC6323986.

18. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003 Nov;13(11):2498-504. doi: 10.1101/gr.1239303. PMID: 14597658; PMCID: PMC403769.

19. Bader GD, Hogue CW. An automated method for finding molecular complexes in large protein interaction networks. BMC Bioinformatics. 2003 Jan 13;4:2. doi: 10.1186/1471-2105-4-2. Epub 2003 Jan 13. PMID: 12525261; PMCID: PMC149346.

20. Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY. cytoHubba: identifying hub objects and sub-networks from complex interactome. BMC Syst Biol. 2014;8 Suppl 4(Suppl 4):S11. doi: 10.1186/1752-0509-8-S4-S11. Epub 2014 Dec 8. PMID: 25521941; PMCID: PMC4290687.

21. Zhou G, Soufan O, Ewald J, Hancock REW, Basu N, Xia J. NetworkAnalyst 3.0: a visual analytics platform for comprehensive gene expression profiling and meta-analysis. Nucleic Acids Res. 2019 Jul 2;47(W1):W234-W241. doi: 10.1093/nar/gkz240. PMID: 30931480; PMCID: PMC6602507.

22. Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. Bioinformatics. 2010 Jun 15;26(12):1572-3. doi: 10.1093/bioinformatics/btq170. Epub 2010 Apr 28. PMID: 20427518; PMCID: PMC2881355.

23. Cox DR. Regression models and life-tables. J R Stat Soc Series B Stat Methodol. 1972;34(2):187-202. doi: 10.1111/j.2517-6161.1972.tb00899.x

24. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008 Dec 29;9:559. doi: 10.1186/1471-2105-9-559. PMID: 19114008; PMCID: PMC2631488.

25. Tibshirani R. The lasso method for variable selection in the Cox model. Stat Med. 1997 Feb 28;16(4):385-95. doi: 10.1002/(sici)1097-0258(19970228)16:4<385::aid-sim380>3.0.co;2-3. PMID: 9044528.

26. Zhang Z, Zhao Y, Canes A, Steinberg D, Lyashevska O; written on behalf of AME Big-Data Clinical Trial Collaborative Group. Predictive analytics with gradient boosting in clinical medicine. Ann Transl Med. 2019 Apr;7(7):152. doi: 10.21037/atm.2019.03.29. PMID: 31157273; PMCID: PMC6511546.

27. Huynh-Cam TT, Chen LS, Le H. Using decision trees and random forest algorithms to predict and determine factors contributing to first-year university students' learning performance. Algorithms. 2021;14(11):318. doi: 10.3390/a14110318.

28. Breiman L. Random forests. Mach Learn. 2001;45(1):5-32. doi: 10.1023/A:1010933404324.

29. Iasonos A, Schrag D, Raj GV, Panageas KS. How to build and interpret a nomogram for can-cer prognosis. J Clin Oncol. 2008;26(8):1364-1370. doi: 10.1200/JCO.2007.12.9791.

30. Harrell FE. Regression modeling strategies. 2nd ed. Springer; 2015. doi: 10.1007/978-3-319-19425-7.

31. Blanche P, Dartigues JF, Jacqmin-Gadda H. Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. Stat Med. 2013 Dec 30;32(30):5381-97. doi: 10.1002/sim.5958. Epub 2013 Sep 12. PMID: 24027076.

32. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making. 2006 Nov-Dec;26(6):565-74. doi: 10.1177/0272989X06295361. PMID: 17099194; PMCID: PMC2577036.

33. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting

genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005 Oct 25;102(43):15545-50. doi: 10.1073/pnas.0506580102. Epub 2005 Sep 30. PMID: 16199517; PMCID: PMC1239896.

34. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdóttir H, Tamayo P, Mesirov JP. Molecular signatures database (MSigDB) 3.0. Bioinformatics. 2011 Jun 15;27(12):1739-40. doi: 10.1093/bioinformatics/btr260. Epub 2011 May 5. PMID: 21546393; PMCID: PMC3106198.

35. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst. 2015 Dec 23;1(6):417-425. doi: 10.1016/j.cels.2015.12.004. PMID: 26771021; PMCID: PMC4707969.

36. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 2000 Jan 1;28(1):27-30. doi: 10.1093/nar/28.1.27. PMID: 10592173; PMCID: PMC102409.

37. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet. 2000 May;25(1):25-9. doi: 10.1038/75556. PMID: 10802651; PMCID: PMC3037419.

38. Yoshihara K, Shahmoradgoli M, Martínez E, Vegesna R, Kim H, Torres-Garcia W, Treviño V, Shen H, Laird PW, Levine DA, Carter SL, Getz G, Stemke-Hale K, Mills GB, Verhaak RG. Inferring tumour purity and stromal and immune cell admixture from expression data. Nat Commun. 2013;4:2612. doi: 10.1038/ncomms3612. PMID: 24113773; PMCID: PMC3826632.

39. Barbie DA, Tamayo P, Boehm JS, Kim SY, Moody SE, Dunn IF, Schinzel AC, Sandy P, Meylan E, Scholl C, Fröhling S, Chan EM, Sos ML, Michel K, Mermel C, Silver SJ, Weir BA, Reiling JH, Sheng Q, Gupta PB, Wadlow RC, Le H, Hoersch S, Wittner BS, Ramaswamy S, Livingston DM, Sabatini DM, Meyerson M, Thomas RK, Lander ES, Mesirov JP, Root DE, Gilliland DG, Jacks T, Hahn WC. Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. Nature. 2009 Nov 5;462(7269):108-12. doi: 10.1038/nature08460. Epub 2009 Oct 21. PMID: 19847166; PMCID: PMC2783335.

40. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M, Alizadeh AA. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods. 2015 May;12(5):453-7. doi: 10.1038/nmeth.3337. Epub 2015 Mar 30. PMID: 25822800; PMCID: PMC4739640.

41. Cancer Genome Atlas Research Network. Comprehensive molecular characterization of urothelial bladder carcinoma. Nature. 2014 Mar 20;507(7492):315-22. doi: 10.1038/nature12965. Epub 2014 Jan 29. PMID: 24476821; PMCID: PMC3962515.

42. Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. Genome Res. 2018 Nov;28(11):1747-1756. doi: 10.1101/gr.239244.118. Epub 2018 Oct 19. PMID: 30341162; PMCID: PMC6211645.

43. Rizvi NA, Hellmann MD, Snyder A, Kvistborg P, Makarov V, Havel JJ, Lee W, Yuan J, Wong P, Ho TS, Miller ML, Rekhtman N, Moreira AL, Ibrahim F, Bruggeman C, Gasmi B, Zappasodi R, Maeda Y, Sander C, Garon EB, Merghoub T, Wolchok JD, Schumacher TN, Chan TA. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. Science. 2015;348(6230):124-128. doi: 10.1126/science.aaa1348.

44. Charoentong P, Finotello F, Angelova M, Mayer C, Efremova M, Rieder D, Hackl H, Trajanoski Z. Pan-cancer Immunogenomic Analyses Reveal Genotype-Immunophenotype Relationships and Predictors of Response to Checkpoint Blockade. Cell Rep. 2017 Jan 3;18(1):248-262. doi: 10.1016/j.celrep.2016.12.019. PMID: 28052254.

45. Jiang P, Gu S, Pan D, Fu J, Sahu A, Hu X, Li Z, Traugh N, Bu X, Li B, Liu J, Freeman GJ, Brown MA, Wucherpfennig KW, Liu XS. Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. Nat Med. 2018 Oct;24(10):1550-1558. doi: 10.1038/s41591-018-0136-1. Epub 2018 Aug 20. PMID: 30127393; PMCID: PMC6487502.

46. Maeser D, Gruener RF, Huang RS. oncoPredict: an R package for predicting in vivo or cancer patient drug response and biomarkers from cell line screening data. Brief Bioinform. 2021 Nov 5;22(6):bbab260. doi: 10.1093/bib/bbab260. PMID: 34260682; PMCID: PMC8574972.

47. Wilcoxon F. Individual comparisons by ranking methods. In: Kotz S, Johnson NL editors. Breakthroughs in Statistics. New York: Springer; 1992:196-202. doi: 10.2307/3001968.

48. Liu CJ, Hu FF, Xia MX, Han L, Zhang Q, Guo AY. GSCALite: a web server for gene set cancer analysis. Bioinformatics. 2018 Nov 1;34(21):3771-3772. doi: 10.1093/bioinformatics/bty411. PMID: 29790900.

49. Wilcoxon F. Individual comparisons by ranking methods. Biometrics Bull. 1945;1(6):80-83. doi: 10.2307/3001968.

50. Spearman C. The proof and measurement of association between two things. Am J Psychol. 1904;15(1):72-101. doi: 10.2307/1412159.

51. Kim SE. Enzymes involved in folate metabolism and its implication for cancer treatment. Nutr Res Pract. 2020;14(2):95-101. doi: 10.4162/ nrp.2020.14.2.95.

52. Zarou MM, Vazquez A, Vignir Helgason G. Folate metabolism: a re-emerging therapeutic target in haematological cancers. Leukemia. 2021 Jun;35(6):1539-1551. doi: 10.1038/s41375-021-01189-2. Epub 2021 Mar 11. PMID: 33707653; PMCID: PMC8179844.

53. Sun W, Liu R, Gao X, Lin Z, Tang H, Cui H, Zhao E. Targeting serine-glycine-one-carbon metabolism as a vulnerability in cancers. Biomark Res. 2023 May 5;11(1):48. doi: 10.1186/s40364-023-00487-4. PMID: 37147729; PMCID: PMC10161514.

54. Shang M, Yang H, Yang R, Chen T, Fu Y, Li Y, Fang X, Zhang K, Zhang J, Li H, Cao X, Gu J, Xiao J, Zhang Q, Liu X, Yu Q, Wang T. The folate cycle enzyme MTHFD2 induces cancer immune evasion through PD-L1 up-regulation. Nat Commun. 2021 Mar 29;12(1):1940. doi: 10.1038/s41467-021-22173-5. PMID: 33782411; PMCID: PMC8007798.

55. Agarwal S, Behring M, Hale K, Al Diffalha S, Wang K, Manne U, Varambally S. MTHFD1L, a folate cycle enzyme, is involved in progression of colorectal cancer. Transl Oncol. 2019;12(11):1461-1467. doi: 10.1016/j.tranon.2019.07.011.

56. He Z, Wang X, Zhang H, Liang B, Zhang J, Zhang Z, Yang Y. High expression of folate cycle enzyme MTHFD1L correlates with poor prognosis and increased proliferation and migration in colorectal cancer. J Cancer. 2020 Apr 25;11(14):4213-4221. doi: 10.7150/jca.35014. PMID: 32368304; PMCID: PMC7196253.

57. Lee D, Xu IM, Chiu DK, Lai RK, Tse AP, Lan Li L, Law CT, Tsang FH, Wei LL, Chan CY, Wong CM, Ng IO, Wong CC. Folate cycle enzyme MTHFD1L confers metabolic advantages in hepatocellular carcinoma. J Clin Invest. 2017 May 1;127(5):1856-1872. doi: 10.1172/JCI90253. Epub 2017 Apr 10. PMID: 28394261; PMCID: PMC5409797.

58. Sial N, Rehman JU, Saeed S, Ahmad M, Hameed Y, Atif M, Rehman A, Asif R, Ahmed H, Hussain MS, Khan MR, Ambreen A, Ambreen A. Integrative analysis reveals methylenetetrahydrofolate dehydrogenase 1-like as an independent shared diagnostic and prognostic biomarker in five different human cancers. Biosci Rep. 2022 Jan 28;42(1):BSR20211783. doi: 10.1042/BSR20211783. PMID: 34908119; PMCID: PMC8738869.

59. Makishima H, Yoshida K, Nguyen N, Przychodzen B, Sanada M, Okuno Y, Ng KP, Gudmundsson KO, Vishwakarma BA, Jerez A, Gomez-Segui I, Takahashi M, Shiraishi Y, Nagata Y, Guinta K, Mori H, Sekeres MA, Chiba K, Tanaka H, Muramatsu H, Sakaguchi H, Paquette RL, McDevitt MA, Kojima S, Sauntharajah Y, Miyano S, Shih LY, Du Y, Ogawa S, Maciejewski JP. Somatic SETBP1 mutations in myeloid malignancies. Nat Genet. 2013 Aug;45(8):942-6. doi: 10.1038/ng.2696. Epub 2013 Jul 7. PMID: 23832012; PMCID: PMC3729750.

60. Díaz MI, Díaz P, Bennett JC, Urra H, Ortiz R, Orellana PC, Hetz C, Quest AFG. Caveolin-1 suppresses tumor formation through the inhibition of the unfolded protein response. Cell Death Dis. 2020 Aug 3;11(8):648. doi: 10.1038/s41419-020-02792-4. PMID: 32811828; PMCID: PMC7434918.

61. Mohiuddin IS, Wei SJ, Kang MH. Role of OCT4 in cancer stem-like cells and chemotherapy resistance. Biochim Biophys Acta Mol Basis Dis. 2020 Apr 1;1866(4):165432. doi: 10.1016/j.bbadis.2019.03.005. Epub 2019 Mar 21. PMID: 30904611; PMCID: PMC6754810.

62. Liu S, Huang H, Lu X, Golinski M, Comesse S, Watt D, Grossman RB, Moscow JA. Down-regulation of thiamine transporter THTR2 gene expression in breast cancer and its association with resistance to apoptosis. Mol Cancer Res. 2003 Jul;1(9):665-73. PMID: 12861052.

63. Zoeller JJ, Whitelock JM, Iozzo RV. Perlecan regulates developmental angiogenesis by modulating the VEGF-VEGFR2 axis. Matrix Biol. 2009 Jun;28(5):284-91. doi: 10.1016/j.matbio.2009.04.010. Epub 2009 May 5. PMID: 19422911; PMCID: PMC2705690.

64. Gajewski TF. The Next Hurdle in Cancer Immunotherapy: Overcoming the Non-T-Cell-Inflamed Tumor Microenvironment. Semin Oncol. 2015 Aug;42(4):663-71. doi: 10.1053/j.seminoncol.2015.05.011. Epub 2015 Jun 3. PMID: 26320069; PMCID: PMC4555998.

65. Ayers M, Lunceford J, Nebozhyn M, Murphy E, Loboda A, Kaufman DR, Albright A, Cheng JD, Kang SP, Shankaran V, Piha-Paul SA, Yearley J, Seiwert TY, Ribas A, McClanahan TK. IFN-γ-related mRNA profile predicts clinical response to PD-1 blockade. J Clin Invest. 2017 Aug 1;127(8):2930-2940. doi: 10.1172/JCI91190. Epub 2017 Jun 26. PMID: 28650338; PMCID: PMC5531419.

66. Mariathasan S, Turley SJ, Nickles D, Castiglioni A, Yuen K, Wang Y, Kadel EE III, Koeppen H, Astarita JL, Cubas R, Jhunjhunwala S, Banchereau R, Yang Y, Guan Y, Chalouni C, Ziai J, Şenbabaoğlu Y, Santoro S, Sheinson D, Hung J, Giltnane JM, Pierce AA, Mesh K, Lianoglou S, Riegler J, Carano RAD, Eriksson P, Höglund M, Somarriba L,

Halligan DL, van der Heijden MS, Loriot Y, Rosenberg JE, Fong L, Mellman I, Chen DS, Green M, Derleth C, Fine GD, Hegde PS, Bourgon R, Powles T. TGFβ attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. Nature. 2018 Feb 22;554(7693):544-548. doi: 10.1038/nature25501. Epub 2018 Feb 14. PMID: 29443960; PMCID: PMC6028240.

67. Robertson AG, Kim J, Al-Ahmadie H, Bellmunt J, Guo G, Cherniack AD, Hinoue T, Laird PW, Hoadley KA, Akbani R, Castro MAA, Gibb EA, Kanchi RS, Gordenin DA, Shukla SA, Sanchez-Vega F, Hansel DE, Czerniak BA, Reuter VE, Su X, de Sa Carvalho B, Chagas VS, Mungall KL, Sadeghi S, Pedamallu CS, Lu Y, Klimczak LJ, Zhang J, Choo C, Ojesina AI, Bullman S, Leraas KM, Lichtenberg TM, Wu CJ, Schultz N, Getz G, Meyerson M, Mills GB, McConkey DJ; TCGA Research Network; Weinstein JN, Kwiatkowski DJ, Lerner SP. Comprehensive Molecular Characterization of Muscle-Invasive Bladder Cancer. Cell. 2017 Oct 19;171(3):540-556.e25. doi: 10.1016/j.cell.2017.09.007. Epub 2017 Oct 5. Erratum in: Cell. 2018 Aug 9;174(4):1033. doi: 10.1016/j.cell.2018.07.036. PMID: 28988769; PMCID: PMC5687509.

68. Li X, Fu S, Huang Y, Luan T, Wang H, Wang J. Identification of a novel metabolism-related gene signature associated with the survival of bladder cancer. BMC Cancer. 2021 Nov 24;21(1):1267. doi: 10.1186/s12885-021-09006-w. PMID: 34819038; PMCID: PMC8611960.

69. Wang Y, Chen L, Yu M, Fang Y, Qian K, Wang G, Ju L, Xiao Y, Wang X. Immune-related signature predicts the prognosis and immunotherapy benefit in bladder cancer. Cancer Med. 2020 Oct;9(20):7729-7741. doi: 10.1002/cam4.3400. Epub 2020 Aug 25. PMID: 32841548; PMCID: PMC7571842.