scientific reports

Check for updates

OPEN

Identification of novel therapeutic targets in hepatitis-B virus-associated membranous nephropathy using scRNA-seq and machine learning

Yongzheng Hu[1], Qian An[2], Xinxin Yu[3] & Wei Jiang[1]✉

Hepatitis B Virus-associated membranous nephropathy (HBV-MN) significantly impacts renal health, particularly in areas with high HBV prevalence. Understanding the molecular mechanisms underlying HBV-MN is crucial for developing effective therapeutic strategies. This study aims to elucidate the roles of miR-223-3p and CRIM1 in HBV-MN using single-cell RNA sequencing (scRNA-seq) and machine learning. scRNA-seq analysis identified a distinct subcluster of podocytes linked to HBV-MN progression. miR-223-3p emerged as a critical regulatory molecule, with overexpression resulting in decreased CRIM1 expression. Dual-luciferase reporter assays confirmed miR-223-3p targeting CRIM1 at a conserved binding site. These findings were corroborated by machine learning models, which highlighted the significance of miR-223-3p and CRIM1 in disease pathology. miR-223-3p plays a pivotal role in modulating CRIM1 expression in HBV-MN, providing a potential therapeutic target. Integrating scRNA-seq with machine learning offers valuable insights into the molecular landscape of HBV-MN, paving the way for novel interventions.

Keywords HBV-MN, MiR-223-3p, CRIM1, ScRNA-seq, Machine learning

Abbreviations

AKI	Acute kidney injury
AUC	Area under the curve
BP	Biological process
CC	Cellular component
CKD	Chronic kidney disease
CTL	Control sample
DEG	Differentially expressed gene
DKD	Diabetic kidney disease
EC	Endothelial cells
GEO	Gene expression omnibus
HBV-MN	Hepatitis b virus-associated membranous nephropathy
hdWGCNA	High dimensional weighted gene co-expression network analysis
IC	Intercalated cells
IMN	Immune cells
KNN	K-nearest neighbors
LDH	Lactate dehydrogenase
LOH	Loop of henle cells
MF	Molecular function

[1]Department of Nephrology, The Affiliated Hospital of Qingdao University, Qingdao, Shandong, China. [2]Department of Nephrology, Qingdao Central Hospital, Qingdao, Shandong, China. [3]Department of Nephrology, Qingdao Eighth People's Hospital, Qingdao, Shandong, China. ✉email: jiangwei866@qdu.edu.cn

miRNAs	MicroRNAs
PC	Principal cells
POD	Podocytes
PT	Proximal tubule cells
RF	Random forest
ROC	Receiver operating characteristic
ROS	Reactive oxygen species
scRNA-seq	Single-cell RNA sequencing
SMC	Smooth muscle cells
XGBoost	Extreme gradient boosting

Hepatitis B virus (HBV) infection remains one of the most significant global public health challenges, affecting nearly 250 million individuals worldwide. Regions such as Africa, Southeast Asia, and the Western Pacific bear a disproportionate burden of chronic HBV infection, which not only predisposes individuals to liver-related complications but also contributes to a range of extrahepatic manifestations, including renal diseases[1]. Among these, HBV-associated membranous nephropathy (HBV-MN) is particularly notable, as it is characterized by the deposition of immune complexes—often containing the hepatitis B surface antigen (HBsAg)—within the glomerular basement membrane[2]. This immune complex deposition initiates local inflammation, ultimately leading to structural and functional alterations in the kidney.

At the core of HBV-MN pathology lies podocyte injury. Damage to these cells results in proteinuria and progressive renal dysfunction, making the elucidation of underlying molecular mechanisms imperative for improving therapeutic strategies. The precise signaling pathways and molecular interactions driving podocyte injury in HBV-MN remain incompletely defined, warranting further investigation[3].

The discovery of microRNAs (miRNAs) is one of the most significant breakthroughs in recent decades. small RNAs can fine-tune inflammatory cascades[4]. Recent advances in molecular biology have highlighted the regulatory roles of microRNAs (miRNAs) in kidney diseases[5]. Among these small non-coding RNAs, miR-223-3p has emerged as a pivotal regulator involved in various renal pathologies, including diabetic kidney disease and chronic kidney disease. Dysregulation of miR-223-3p has been linked to altered cellular processes such as inflammation, apoptosis, and repair mechanisms[6]. However, its specific contribution to the pathogenesis of HBV-MN and the precise mechanisms by which it regulates critical target genes remain to be fully elucidated.

Parallel to these developments, the advent of single-cell RNA sequencing (scRNA-seq) has revolutionized our ability to explore the transcriptomic landscape of complex tissues at unprecedented resolution[7]. The integration of machine learning techniques offers a robust framework for analyzing high-dimensional genomic data[8]. When combined with scRNA-seq data, these algorithms can enhance the precision of gene target identification and facilitate the discovery of intricate molecular interactions underlying HBV-MN.

In this study, we aim to bridge existing knowledge gaps by investigating the role of miR-223-3p in HBV-MN through an integrated analysis combining scRNA-seq with advanced machine learning methodologies. Ultimately, these findings may pave the way for the development of novel, targeted therapeutic interventions to mitigate renal dysfunction associated with HBV infection.

Materials and methods

Cell culture and transfection

Human podocytes were acquired from BeNa Culture Collection (Henan, China) and cultured in RPMI 1640 medium (Pronoza, Cat.No PM150110) supplemented with 10% fetal bovine serum (GIBCO, Cat.No 10099141) and 1% penicillin/streptomycin solution (Meilune, Cat.No MAO110), in a CO2 incubator (PHCBI, MCO-18 AC) at 37 °C with 5% CO2. The plasmid encoding the HBx gene (pHBx) and the corresponding negative control plasmid were constructed and purified by Biogenetech (Anhui, China), ensuring high purity and concentration suitable for transfection. Podocytes were seeded onto 6-well plates at a density of 5×10^4 cells/cm^2 to allow for optimal growth and confluency. Cells were grown until they reached approximately 70% confluence, which is the optimal condition for transfection. Transfection was carried out using Lipofectamine 3000 (Invitrogen, USA) according to the manufacturer's instructions. The experiment included four groups: control group: cells with no treatment; empty plasmid group: cells transfected with empty plasmid; HBx group: cells transfected with the pHBx plasmid; HBx + miR-223-3p mimic group: cells co-transfected with pHBx and miR-223-3p mimic to study the interactive effects on cell behavior and gene regulation. Transfection efficiency was assessed using quantitative polymerase chain reaction (qPCR) to detect the expression levels of specific genes associated with the plasmid vectors and miR-223-3p mimic.

ROS level determination

Podocytes were washed with PBS and incubated with 10 μM DCFDA from the ROS ELISA Kit (Shanghai Kexing, Cat.No F9689-B) for 30 min at 37 °C in darkness. Post-incubation, cells were washed to remove excess DCFDA, lysed, and centrifuged at 12,000×*g* for 15 min at 4 °C (Neofuge 13R, Heal Force). 100 μl of supernatants were added to a 96-well plate and processed as per the kit's instructions. Fluorescence was measured at 450 nm using a Multiskan Sky microplate reader (Thermo Fisher). The optical density(OD) values obtained were normalized to the control group to assess ROS levels.

LDH level determination

Podocytes underwent a PBS wash and subsequent lysis. Following this, cell lysates were subjected to centrifugation at 12,000×*g* for 15 min at a temperature of 4 °C using a Neofuge 13R centrifuge (Heal Force). The

supernatants were then carefully collected, with 100 µl transferred into each well of a 96-well plate, prepared in accordance with the protocol provided with the LDH ELISA Kit (Shanghai Kexing, Cat.No F0350-B). Absorbance levels were then assessed at 450 nm using the Multiskan Sky microplate reader (Thermo Fisher). To gauge cellular damage, the OD measurements were standardized against the control group, focusing on the LDH levels indicated.

Cell viability assessment using MTT Assay

Podocytes were trypsinized using 0.25% trypsin (Biosharp, Cat.No PGY0021) and were then seeded in a 96-well plate (JET Biofil, Cat.No TCP011096) at a density of 5,000 cells per well and allowed to adhere overnight. After treatment according to the experimental groups, 20 µl of MTT solution (Sangon Biotech, Cat.No CB0331) was added to each well and incubated for 4 h at 37 °C. Post-incubation, the medium was removed, and 150 µl of dimethyl sulfoxide (DMSO) was added to dissolve the formazan crystals. The plate was gently shaken for 10 min to ensure complete dissolution. The absorbance was measured at 490 nm using a Multiskan Sky microplate reader (Thermo Fisher). OD values obtained were used to calculate the cell proliferation rate across different treatment groups. Results were expressed as OD490/fold, which represents the proliferation fold increase from 24 to 96 h relative to the 24-h measurement for each group.

Quantitative real-time PCR (qPCR)

48 h post-transfection, cells were lysed using a high-speed low-temperature tissue grinder (Servicebio, KZ-III-F). RNA was extracted using TRIzon Reagent (Kangwei Century, CW0580S) and purified using UltraPure RNA Kit (Kangwei Century, CW0581M), followed by centrifugation in a tabletop high-speed refrigerated centrifuge (Hettichi, 0004219-11). Reverse transcription was performed with Evo M-MLV RT Premix (Acrobiosystems, AG11706) inside a clean bench (Suzhou Purification, SW-CJ-1 FD). The qPCR was conducted on a CFX Connect Real-Time PCR system (Bio-Rad, CFX connect), using SYBR Green Pro Taq HS qPCR Kit (Acrobiosystems, AG11701). The reaction consisted of 12.5 µL SYBR Green mix, 1 µL each of forward and reverse primers (Beijing Dingguo Biotechnology), and cDNA in a total volume of 25 µL. Cycling conditions were 95 °C for 2 min, followed by 40 cycles at 95 °C for 3 s, 60 °C for 30 s, and 70 °C for 10 s. A melting curve was analyzed post-amplification. The internal reference for miRNA-223-3p was U6, and for other mRNAs it was GAPDH. The primer sequences are shown in Table 1. Gene expression was quantified using the 2^-ΔΔCT method and analyzed on the mµLtiskan Sky microplate reader (Thermo Fisher). Primer sequence information for the PCR experiments is provided in Table 1. Each experiment was performed in triplicate.

Western blot

Cells were lysed with enhanced RIPA buffer (Biosharp, AR0102) containing 100 mM PMSF (Biosharp, AR1192). Protein concentrations were determined using a BCA Kit (Beyotime, P0012). Lysates were mixed with 5 × loading buffer (Biosharp, AR1112-10) and heated for denaturation. Proteins were separated by 12.5% SDS-PAGE (Yeasen, PG113) and transferred to PVDF membranes, 0.45 µm for high (Millipore, IPVH00010) and 0.22 µm for low molecular weight proteins (Millipore, ISEQ00010). Membranes were blocked with 5% milk in TBS with 0.1% Tween 20 (Solarbio, T8220) and incubated overnight at 4 °C with primary antibodies against CRIM1 (Thermo Fisher, PA5-34410, 1:1000) and GAPDH (Proteintech, 60004–1-Ig, 1:10000). After washing, they were incubated with HRP-conjugated secondary antibodies (Zhongshan Golden Bridge, ZB-2305 for mouse, ZB-2301 for rabbit, 1:10000). Chemiluminescent detection was performed using an ECL kit (Servicebio, G2014). Band intensity was quantified using ImageJ, with GAPDH as the loading control.

Dual-luciferase reporter (DLR) assay

293 T cells were revived and cultured under standard conditions. For co-transfection, 293 T cells were seeded in 96-well plates at 2 × 10^4 cells/well in DMEM with 10% FBS. The next day, cells were co-transfected with plasmids containing the CRIM1 3'UTR and miR-223-3p using Lipofectamine 2000. After 48 h, luciferase activity

Real time-PCR primer sequences		
Gene	**Sequence**	**5'–3'**
HBx	Forward	AGCAATGTCAACGACCGACCTTG
	Reverse	GACCAATTTATGCCTACAGCCTCCTAG
mir-223-3p	Forward	TGTCAGTTTGTCAAA
	Reverse	CAGTGCGTGTCGTGGAGT
CRIM1	Forward	TTAGGTGTGCGTGGTGCATC
	Reverse	CCGAGGCAATGGGTTTCAAC
GAPDH	Forward	TGTGGGCATCAATGGATTTGG
	Reverse	ACACCATGTATTCCGGGTCAAT
U6	Forward	GCTCGCTTCGGCAGCACA
	Reverse	GAACGCTTCACGAATTTGCGTG

Table 1. Primer sequences used for quantitative real-time PCR (qPCR).

was measured using the Dual-Luciferase Reporter Assay System (Promega). Luminescence was recorded by first adding LAR II reagent, followed by Stop & Glo® Reagent, and reading the signal on a luminometer.

Data collection

The total of 2 single-cell transcriptomics datasets, including 1 HBV-MN samples and 2 health sample, were collected from public dataset Gene Expression Omnibus (GEO). The GSE199850 dataset contained one sample from human kidney from patient with HBV-associated MN[9]. The GSE171458 dataset analyzed kidney samples from 6 patients with IMN and 2 healthy control subjects using single-cell RNA sequencing, healthy control samples were selected by our study[10]. All the datasets were processed in R (V.4.2.3), and the results were showed using ggplot2 R package (V.3.4.3) except where mentioned.

Data quality control and preprocessing

Quality control was rigorously applied following the generation of the expression matrix using Seurat V4.4.0[11].. For downstream analysis, we selected cells with unique molecular identifiers (UMIs) ranging from 0 to 30,000. Cells with fewer than 200 or more than 5,000 unique expressed genes were excluded to ensure high data quality. Doublet cells were identified and removed using DoubletFinder, enhancing the accuracy of differential gene expression analysis. Additionally, cells were excluded if over 50% of their transcriptomes were comprised of mitochondrial genes, which is indicative of cellular stress or damage. After these stringent quality control measures and further corrections for batch effects, a total of 12,521 cells were analyzed, including 8,982 from healthy controls and 3,539 from HBV-MN patients. The data underwent natural log transformation and normalization to facilitate robust downstream analyses.

Cell type classification and marker genes analysis

The Seurat package was utilized to analyze RNA-Sequencing data, facilitating cell type identification and clustering analysis. We employed the shared nearest neighbor (SNN) model for clustering and visualized cellular distribution through dimension reduction techniques including PCA, tSNE, and UMAP. Differences among clusters were assessed using the Wilcoxon rank-sum test, and marker genes for each cluster were identified through the FindAllMarkers function, which implements a likelihood-ratio test to refine the differential gene list. Further, sub-clustering analysis of podocytes was conducted using Seurat's SubsetData function. This comprehensive approach enabled a nuanced exploration of the cellular heterogeneity within the dataset.

Differentially expressed genes analysis

To identify potential differentially expressed genes (DEGs), we employed the Seurat FindMarkers function, which utilizes the Wilcoxon likelihood-ratio test with default parameters. Criteria for selecting DEGs included expression in over 10% of cells within a cluster, an average log-fold change greater than 0.25, and a p-value less than 0.05. Additionally, we characterized the DEGs by comparing the transcriptional profile of HBV-MN with that of healthy donor subjects, enabling a focused investigation into the molecular distinctions driving pathological changes.

Cell type annotation

Cell types were initially annotated automatically using SingleR(V.2.0.0), followed by manual verification based on canonical marker genes[12]. The top DEGs in each cluster were determined by the expression patterns of these canonical markers to accurately define cell types. Seurat was utilized to visualize the expression of these markers through heatmaps and dot plots. Additionally, cells identified as doublets, expressing markers for multiple cell types, were manually excluded from the analysis to ensure the accuracy of the cell type classification.

High dimensional WGCNA (hdWGCNA)

In this study, we applied the hdWGCNA R package for weighted gene co-expression network analysis (WGCNA) of scRNA-seq data, focusing on podocytes identified by Seurat[13]. We filtered genes expressed in fewer than 5% of cells and normalized data using the 'sctransform' function. Pearson correlations were calculated to assess gene interconnectivity. Using the 'pickSoftThreshold' function, we established a scale-free network, generating an adjacency and topological overlap matrix (TOM). Co-expression modules were delineated using the 'dynamicTreeCut' algorithm, assigning each a unique color. Hub genes within modules were evaluated for connectivity using eigengene-based connectivity (kME) values from the 'ModuleConnectivity' function. High kME values identified central genes to each module. Module eigengenes (MEs) were computed with the 'GetMEs' function, summarizing gene expression profiles. The top 25 hub genes, identified by their kME values through the 'GetHubGenes' function, were selected for further analysis, including functional enrichment and clinical correlations[14]. This streamlined approach with hdWGCNA version 0.2.14 ensured accurate and reproducible findings.

Trajectory analysis

To unravel the dynamic gene expression changes across single-cell pseudotime trajectories, we employed the Monocle 2 (version 2.26.0) specifically for podocyte subtypes[15]. Monocle's single-cell trajectory analysis technique uses an algorithm to map the sequence of gene expression changes each cell undergoes during a biological process, placing each cell at its respective position along the trajectory. This method allows for a comprehensive visualization of cellular transitions in gene expression. Criteria for selecting genes to order cells included a presence in at least 10 cells, a mean expression value of at least 0.05, and a dispersion empirical value of at least 2. Cells were subsequently ordered along this trajectory, and their progression was visualized in a reduced dimensional

space to illustrate the developmental continuum. Significant changes in gene expression along the pseudotime trajectory were identified using Monocle's differential Gene Test function, applying a stringent q-value cutoff of less than 0.01. These changes were depicted through branched-heatmap and scatter plot images, providing a detailed representation of gene expression dynamics that define cell state transitions within podocytes.

Cell–cell interaction analysis

After cell type identification, we explored cell–cell communication using CellChat (version 1.6.1)[16]. Initially, a new CellChat object was established based on the merged Seurat object. We utilized the paracrine/autocrine signaling interaction dataset from CellChatDB as the reference database for analysis. Communication probabilities were calculated using the 'computeCommunProb' function with the"truncatedMean"type, setting a trimming parameter of 20% (trim = 0.2). This step was crucial for accurately estimating the likelihood of communication based on a robust statistical approach. Subsequently, cell–cell communication was inferred, and the network was aggregated using the default settings within CellChat. This process synthesized the individual communication events into a comprehensive network, enhancing the interpretability of the data. The aggregation and analysis culminated in the visualization of the number of interactions, which illustrated the extensive cell–cell communication network.

Prediction of target genes

To predict the downstream target genes of miR-223-3p, we utilized three established databases: miRDB, miRnet, and TargetScan[17–19]. These databases are instrumental for identifying potential regulatory relationships based on existing biological data. To enhance the reliability of our predictions, we applied the VennDiagram package(V.1.7.3) in R to determine the intersections among the gene sets predicted by each database. This intersection approach helped refine our results, ensuring that only the most likely target genes, consistently identified across multiple predictive platforms, were considered in our analysis.

Machine learning

To prioritize the most significant genes, we employed three machine learning techniques and divided the data into a testing set (70%) and a training set (30%). To safeguard against overfitting, each model was subjected to fivefold cross-validation. First, we utilized the RandomForest(RF) algorithm(V.4.7.1.1), employing the randomForest package[20]. RF is an ensemble learning method that constructs a multitude of decision trees, each trained on a random subset of features. Importantly, RF calculates the average contribution of each feature, which allows us to rank genes based on their importance in the model—making it particularly suitable for our high-dimensional data. By assessing the contribution values, we ranked the genes based on their importance in the model. Next, we used the xgboost package (V.1.7.5.1) for implementing extreme gradient boosting (XGBoost). XGBoost is a powerful integrated learning algorithm known for its scalability and efficiency in handling diverse data types, enhancing both model visualization and optimization[21]. It uses a gradient boosting framework to iteratively refine the model by minimizing the prediction error, thereby capturing intricate interactions among features. This iterative optimization not only enhances model performance but also provides a clear visualization of feature importance, complementing the insights gained from RF. Lastly, we applied the k-nearest neighbors (KNN) algorithm using the kknn package (V.1.3.1). KNN is a supervised learning algorithm that classifies data points based on the proximity of similar data in the training set, utilizing a majority voting rule among the k-nearest neighbors to determine the classification[22]. Although KNN is less complex than the ensemble methods, it serves as an independent validation tool, ensuring that the gene ranking and classification achieved by RF and XGBoost are robust and consistent. To evaluate the diagnostic performance of each model, we analyzed the receiver operating characteristic (ROC) and calculated the area under the curve (AUC).

Statistical analysis

All data visualization and statistical analyses were performed using R software. We assessed differences between two groups for continuous variables using the independent t-test and for categorical variables using the Chi-squared (X^2) test. A two-sided p-value of less than 0.05 was deemed to indicate statistical significance. This approach ensured a rigorous evaluation of the data, facilitating robust conclusions from our study.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the authors used ChatGPT to assist in drafting and refining the manuscript content. After using this tool, the authors reviewed and edited the content as needed and takes full responsibility for the content of the publication.

Results

Protective role of miR-223-3p in HBV-MN

To investigate the impact of miR-223-3p on podocytes within the context of HBV-MN, we initially engineered a human renal podocyte line to incorporate a plasmid-mediated HBx gene transfection. This was further corroborated by PCR assays, which confirmed the establishment of the HBV-MN model (Fig. 1A). Following this, we assessed miR-223 expression across various podocyte cohorts. Our findings elucidated that HBx exposure markedly diminished miR-223 levels in podocytes (Fig. 1B), a phenomenon paralleled by reduced cellular viability and augmented cell injury (Fig. 1C,D). Subsequent transfection of miR-223 into the HBx-modified podocyte model revealed that miR-223 overexpression mitigates cell damage and enhances cell survival (Fig. 1E). These

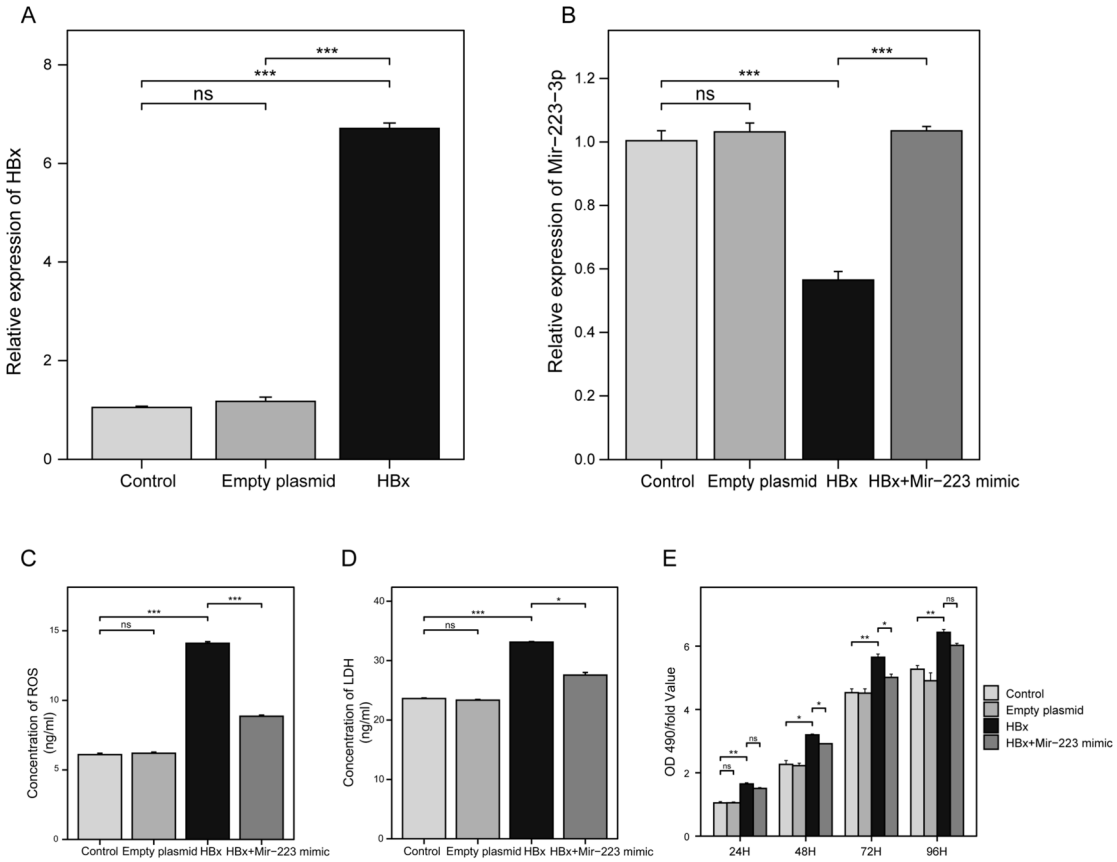

Fig. 1. Effects of HBx transfection and miR-223-3p on podocyte injury and viability in an in vitro HBV-MN model. (**A**) HBx gene expression in human renal podocytes confirms successful establishment of the HBV-MN model. (**B**) miR-223-3p expression in podocytes showing a significant decrease upon HBx transfection and partial rescue by miR-223-3p mimic. (**C**) ROS levels indicate cellular stress in podocytes post-HBx exposure. (**D**) LDH release as a marker of cell injury in HBx-affected podocytes. (**E**) MTT assay illustrating cell viability across different time points (24–96 h) and treatment groups. HBV-MN: HBV-induced membranous nephropathy; ROS: reactive oxygen species; LDH: lactate dehydrogenase. (ns: no significance, *p < 0.05, **p < 0.01, ***p < 0.001.).

observations underscore the potential of miR-223 as a protective agent in podocytes against HBV-MN-induced pathology.

Single-cell sequence in HBV-MN and healthy kidneys

In our investigation, we meticulously analyzed the transcriptomic profiles of 14,340 kidney cells sourced from two individuals without disease and one patient with HBV-MN. Following rigorous processing and stringent filtering of the raw data, we retained 12,521 cells for in-depth analysis. Post-normalization of gene expression levels and conducting principal component analysis (PCA), graph-based clustering facilitated the segmentation of these cells into 10 distinct clusters (Fig. 2A). Notably, the distribution of cell lineages exhibited significant variation across individuals (Fig. 2B). Known kidney cell marker gene expression patterns were subsequently analyzed to accurately define each cell cluster (Table 2).Through the identification of marker genes and the analysis of the top five differentially expressed genes, these clusters were accurately categorized into eight recognized cell lineages (Figs. 2C,D). This analysis highlights the pronounced cellular heterogeneity present in HBV-MN compared to healthy renal tissue, offering profound insights into the pathophysiological underpinnings of the disease.

Co-expression networks in podocytes

Subsequently, we embarked on a high-definition weighted gene co-expression network analysis (hdWGCNA) focusing on podocyte clusters, as depicted in Fig. 3A. Utilizing the TestSoftPowers function (with networkType set to'signed'), we meticulously scanned a range of soft power thresholds (from 1 to 30), facilitating the identification and merging of module similarities. An optimal soft threshold of 12 was determined, enabling the construction of a co-expression matrix tailored for the single-cell transcriptome (with a fraction parameter set to 0.05) (Fig. 3B). This approach culminated in the delineation of two distinct modules, illustrated in Figs. 3C,D. The comprehensive enrichment analysis of module genes within the podocyte cluster is presented in Fig. 3E. Kyoto encyclopedia of genes and genomes(KEGG) enrichment analyses were conducted to explore functional pathways[23–25]. Core genes were subsequently isolated for KEGG pathway enrichment analysis, the findings of which are conveyed in Fig. 3F. A module differential analysis across various groups revealed that disparities in

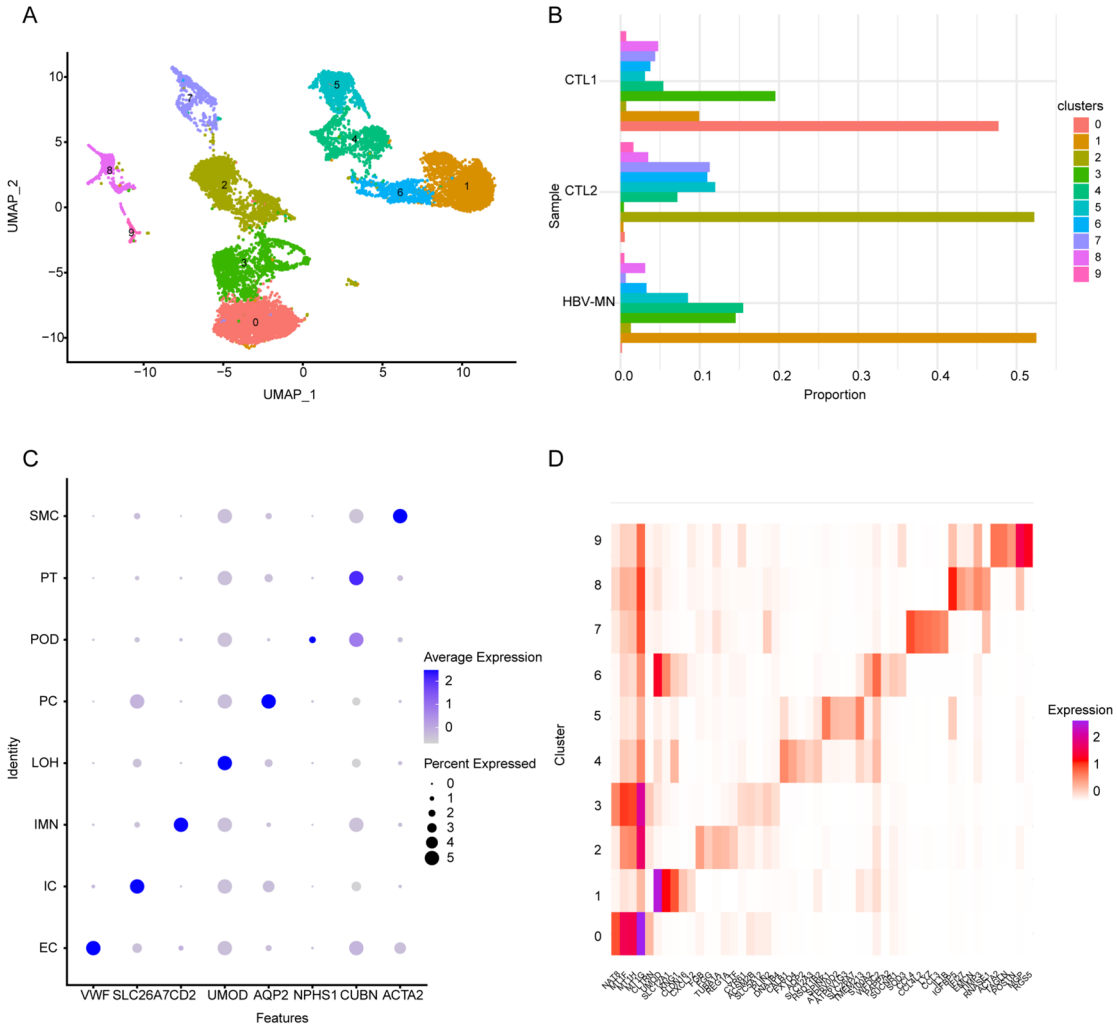

Fig. 2. Transcriptomic profiling and cellular heterogeneity in HBV-MN and healthy kidneys. (**A**) UMAP plot showing 12,521 cells grouped into distinct clusters based on scRNA-seq data. Each color corresponds to a unique cluster. (**B**) Bar plot comparing the proportion of each cluster between healthy control samples (CTL1, CTL2) and the HBV-MN sample, highlighting shifts in cell lineage distribution. (**C**) Dot plot displaying the expression of canonical marker genes (color intensity) and the proportion of cells expressing each marker (dot size) across clusters, facilitating cell type annotation. (**D**) Heatmap of the top 5 most differentially expressed genes in each cluster. *HBV-MN* HBV-induced membranous nephropathy; *CTL* control sample; *SMC* smooth muscle cells; *PT* proximal tubule cells, *POD* podocytes, *PC* principal cells, *LOH* loop of Henle cells, *IMN* immune cells, *IC* intercalated cells, *EC* endothelial cells.

Markers for different cell types		
Marker	**Celltype**	**References**
VWF	Endothelial cells	Vijay et al.[28]
SLC26 A7	Intercalated cells	Wu et al.[29]
CD2	Immune cells	Izar et al.[30]
UMOD	Loop of Henle cells	Wu et al.[29]
AQP2	Principal cells	Wilson et al.[31]
NPHS1	Podocytes	Czerniecki et al.[32]
CUBN	Proximal tubule cells	Wu et al.[29]
ACTA2	Smooth muscle cells	Zhang et al.[33]

Table 2. Markers for different kidney cell types used for cell type identification.

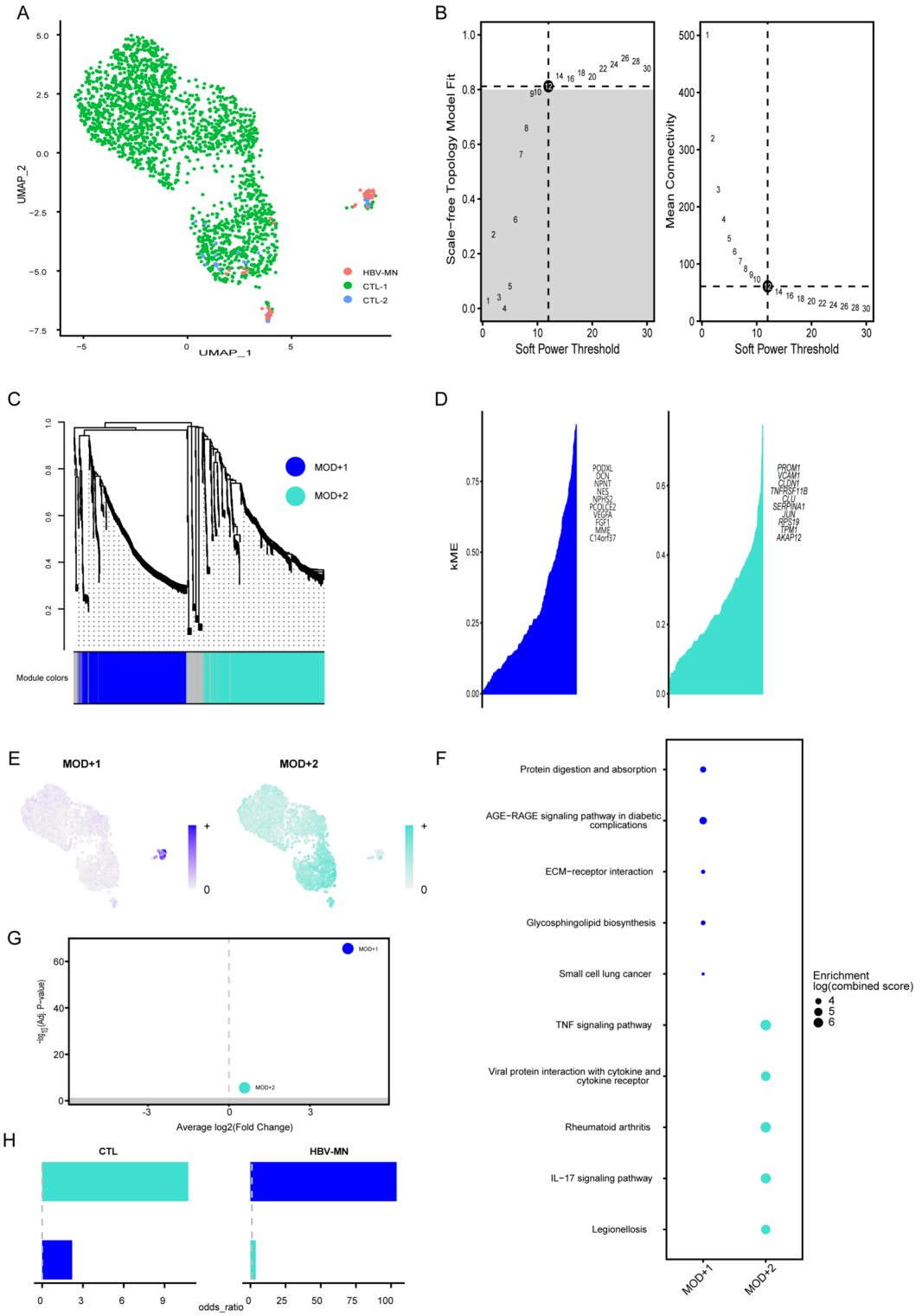

Fig. 3. Identifcation of gene co-expression modules among podocytes. (**A**) UMAP plot showing the distribution of podocyte subpopulations in healthy control (CTL) and HBV-MN samples, illustrating distinct clusters within the podocyte compartment. (**B**) Soft power threshold optimization for constructing a scale-free network in hdWGCNA. The left panel displays the scale-free topology fit index across different power values, while the right panel shows mean connectivity, aiding in the selection of an optimal soft power. (**C**) Hierarchical clustering dendrogram depicting gene co-expression modules in podocytes. Modules are color-coded, reflecting groups of highly correlated genes. (**D**) Top 10 eigengenes for each module, ranked by eigengene-based connectivity (kME). Genes with higher kME values are considered central or "hub" genes within their respective modules. (**E**) UMAP visualization of module expression patterns among podocyte clusters, highlighting spatial differences in module-specific gene expression. (**F**) Top 5 KEGG pathway enrichment analysis of each module. (**G**) Module differential analysis across various groups. (**H**) Overlap of characteristic genes between significant modules and each disease group. *HBV-MN* HBV-induced membranous nephropathy, *CTL* control sample; *MOD* module.

Module 1 were markedly pronounced (Fig. 3G), with a significant overlap of characteristic genes between Module 1 and the HBV-MN group, detailed in Fig. 3H. This analysis underscores the intricate co-expression networks within podocytes, highlighting potential molecular pathways implicated in HBV-MN.

Podocyte differentiation dynamics

Our study extends to a pseudotime analysis, capturing the developmental and differentiation trajectory of podocytes across varied groups, revealing a critical branching point leading to three distinct branches (Fig. 4A). Through integrating biological relevance with statistical methodologies, we established the direction of the pseudotime differentiation trajectory (Fig. 4B). This analysis unveiled no significant stage disparities in differentiation among healthy podocytes. Conversely, podocytes from the HBV-MN group predominantly converged at the terminal phase of the differentiation spectrum, specifically Branch 3 (Fig. 4C,D). Further exploration through pseudotime trajectory analysis yielded the top 50 differentially expressed genes across the branches, shown in Fig. 4E. Beam analysis facilitated the identification of 50 pivotal genes instrumental in cell state transitions, exhibiting progressive upregulation or downregulation during cell differentiation and categorized into four clusters (Fig. 4F). Notably, genes within Cluster 3 demonstrated heightened expression in Branch 3, closely aligning with the HBV-MN podocyte phenotype along the pseudotime axis. Subsequent functional clustering and GO enrichment analysis indicated that genes in Branch 3 predominantly contribute to kidney development, the composition of collagen-containing extracellular matrices, and insulin-like growth factor binding, as highlighted in Fig. 4G. This comprehensive pseudotime analysis elucidates the intricate differentiation patterns of podocytes, with particular emphasis on the altered trajectory observed in HBV-MN, offering valuable insights into the molecular mechanisms underpinning disease progression.

Inter-cellular communication in HBV-MN

To elucidate the intercellular dynamics within HBV-MN and compare them with those in healthy tissue, we employed CellChat for an in-depth assessment of cell interactions, encompassing both the volume and intensity of cellular communications. By aggregating communication probabilities across the informational network, we were able to quantify the disparities in overall information flow between HBV-MN afflicted and normal tissues. Our analysis revealed a diminution in both the aggregate number of signals and their intensity within the HBV-MN cohort relative to the control group (Fig. 5A). Notably, signaling pathways such as CD46, ANNEXIN, and VASFATIN were predominantly active in normal tissues, whereas PTN, VEGF, and DESMOSOME pathways were more prevalent in HBV-MN tissues. Further scrutinizing ligand-receptor interactions between podocytes and other cell types revealed differential patterns between HBV-MN and healthy counterparts (Fig. 5B,C). Specifically, podocytes in HBV-MN exhibited augmented communications with most cell types, particularly exhibiting intensified interactions with endothelial cells, albeit with attenuated communications with proximal tubular cells, in contrast to normal tissues. Delving deeper, we identified pivotal ligand-receptor pairs in HBV-MN podocytes; notably, the CD99-CD99 and several VEGFA and COL4 A3 receptor pairs were enhanced when podocytes served as the signaling origin. Conversely, when receiving signals, the affinity of FN1, COL4 A1, and COL1 A2 related receptor pairs markedly declined. Intriguingly, the APP-CD74 receptor pair showcased increased communication strength when podocytes were the signal initiators but exhibited reduced interaction when they assumed the receptor role (Fig. 5D). This comprehensive analysis underscores the nuanced alterations in cellular communication networks within HBV-MN, shedding light on the complex interplay of signaling pathways that may underlie disease pathology.

Identifying miR-223 targets in HBV-MN

Leveraging three distinguished online databases, we embarked on a quest to unveil potential target genes of miR-223, uncovering that miRDB, mirnet, and TargetScan forecasted 519, 168, and 412 candidate genes, respectively, as detailed in the Supplementary Material. Integrating these predictions with our single-cell sequencing data, we employed a Venn diagram to intersect the paramount 30 genes from Module 1—aligned with the HBV-MN expression pattern discerned through hdWGCNA analysis—and the leading 30 genes identified by BEAM analysis in pseudotime analysis as pivotal in directing differentiation towards HBV-MN (notably, Branch 3). This confluence elucidated CRIM1, MYL9, and TMSB4X as the intersecting genes, showed in Fig. 6A. A further intersection of these genes with the anticipated targets proposed by the databases highlighted CRIM1 and MYL9 as the definitive targets of miR-223 (Fig. 6B). To refine our selection of candidate genes, we applied sophisticated machine learning algorithms—XGBoost, RF, and KNN. Initially, we randomly allocated 70% of our single-cell sequencing samples to serve as a training set, reserving the remainder for validation. The ROC curves demonstrated the algorithms'proficient predictive capabilities within the training ensemble (Fig. 6C). Subsequently, genes were appraised based on their calculated significance, with CRIM1 consistently outperforming MYL9 in importance rankings (Fig. 6D&5E). Further analysis employing the KNN algorithm revealed nearly identical predictive accuracy for both genes (Fig. 6F). Consistent with our initial findings, the ROC analysis in the validation set reaffirmed CRIM1's superior predictive prowess (Fig. 6G). This meticulous approach to gene target identification underscores the synergy between bioinformatics tools and machine learning in elucidating the molecular underpinnings of HBV-MN, with CRIM1 emerging as a prominent target of miR-223.

MiR-223-3p mediated regulation of CRIM1 expression

To validate the interaction between miR-223-3p and CRIM1 inferred from our analyses, we assessed CRIM1 expression levels in various experimental conditions. Our results confirmed an upregulation of CRIM1 expression in the overexpressed HBx group, corroborating our single-cell sequencing data (Fig. 7A). Introduction of overexpressed miR-223-3p significantly reduced CRIM1 expression, suggesting miR-223-3p's suppressive role

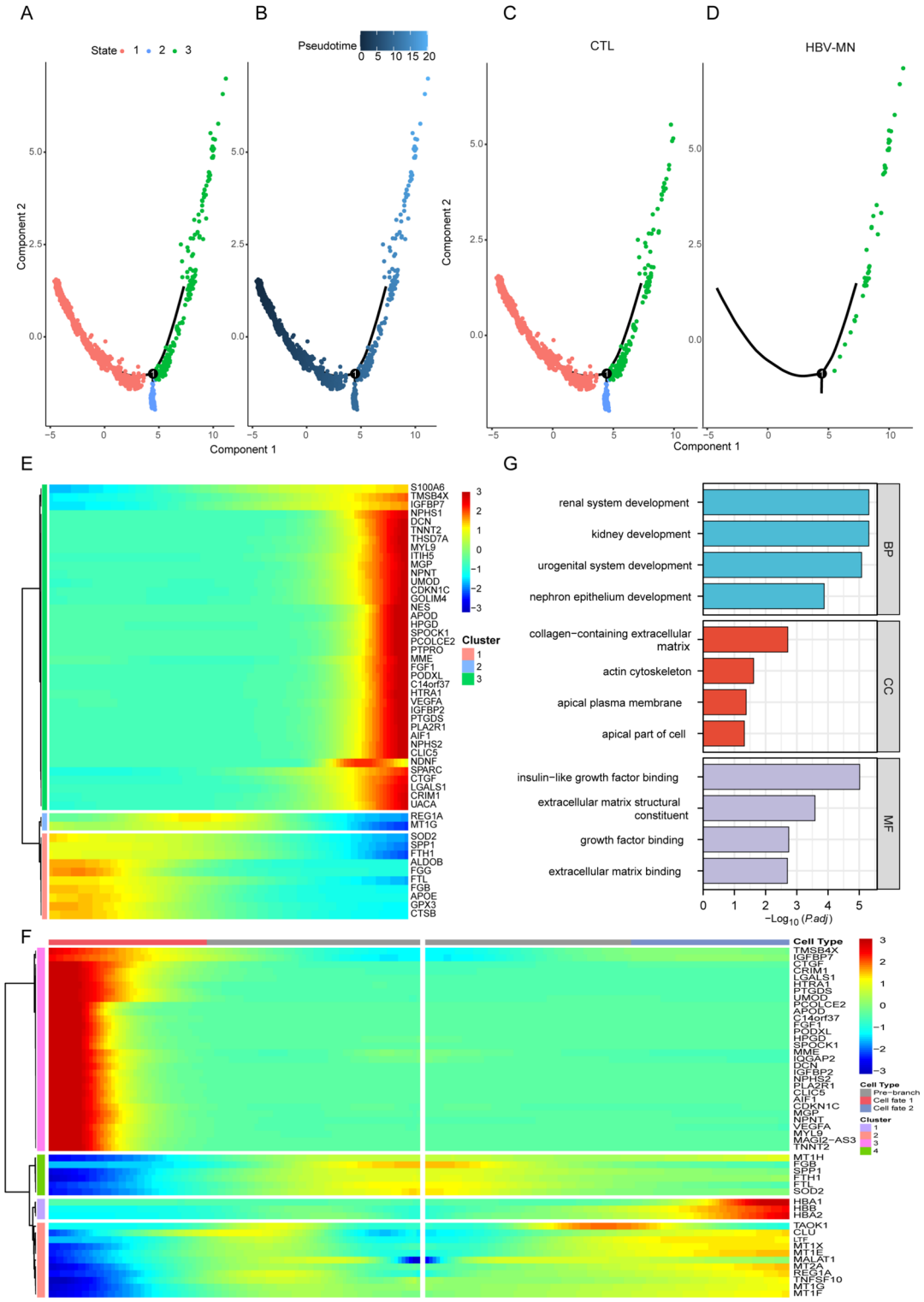

Fig. 4. Pseudotime analysis revealing podocyte differentiation dynamics. (**A**) Pseudotime trajectory delineating three distinct podocyte states (labeled 1, 2, and 3), illustrating the progression of cellular differentiation. (**B**) Directionality of podocyte differentiation with pseudotime trajectory analysis. (**C**,**D**) Differential podocyte differentiation in healthy vs. HBV-MN conditions. (**E**) Top 50 differentially expressed genes across differentiation stages. (**F**) Beam analysis revealed key genes driving podocyte evolution. (**G**) GO enrichment analysis of genes in terminal differentiation of podocytes. *HBV-MN* HBV-induced membranous nephropathy, *CTL* control sample, *BP* biological process, *CC* cellular component, *MF* molecular function.

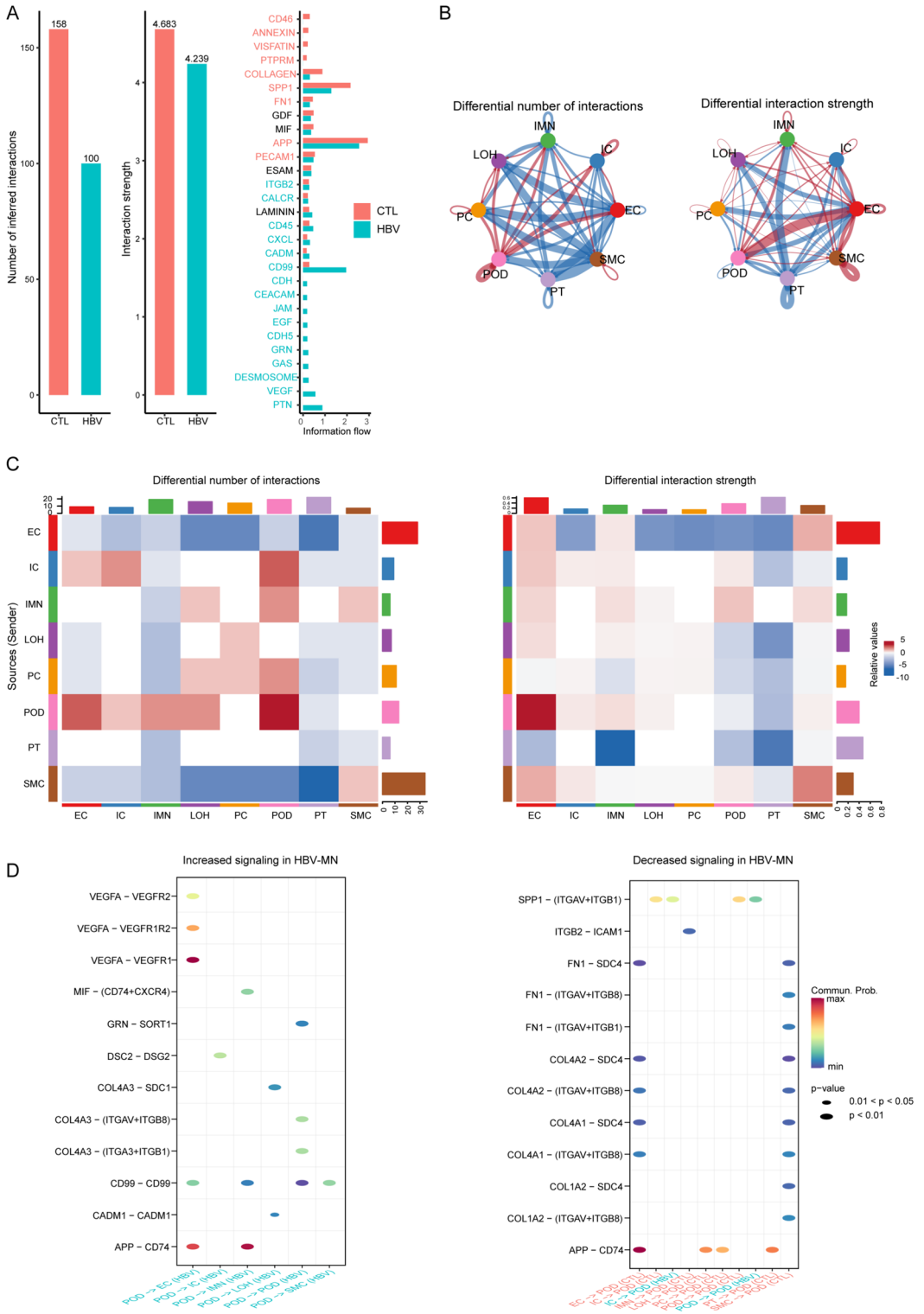

Fig. 5. Intercellular communication through CellChat analysis. (**A**) Overview of total signaling interactions and overall communication strength comparing HBV-MN (blue) and healthy kidney tissues (red). Bar plots indicate both the number of signals and their cumulative intensities. (**B**) Comparative chord diagram of podocyte interactions in HBV-MN vs. healthy conditions. Each node represents a cell type, and edge thickness corresponds to the strength or number of the interaction. (**C**) Heatmap analysis of podocyte communication patterns. Attenuated/downregulated in blue and enhanced/upregulated in red. (**D**) Increased and decreased key ligand-receptor interactions. *HBV-MN* HBV-induced membranous nephropathy, *CTL* control sample, *SMC* smooth muscle cell; *PT* proximal tubule cells; *POD* podocytes; *PC* principal cells; *LOH* loop of Henle cells, *IMN* immune cells; *IC* intercalated cells; *EC* endothelial cells.

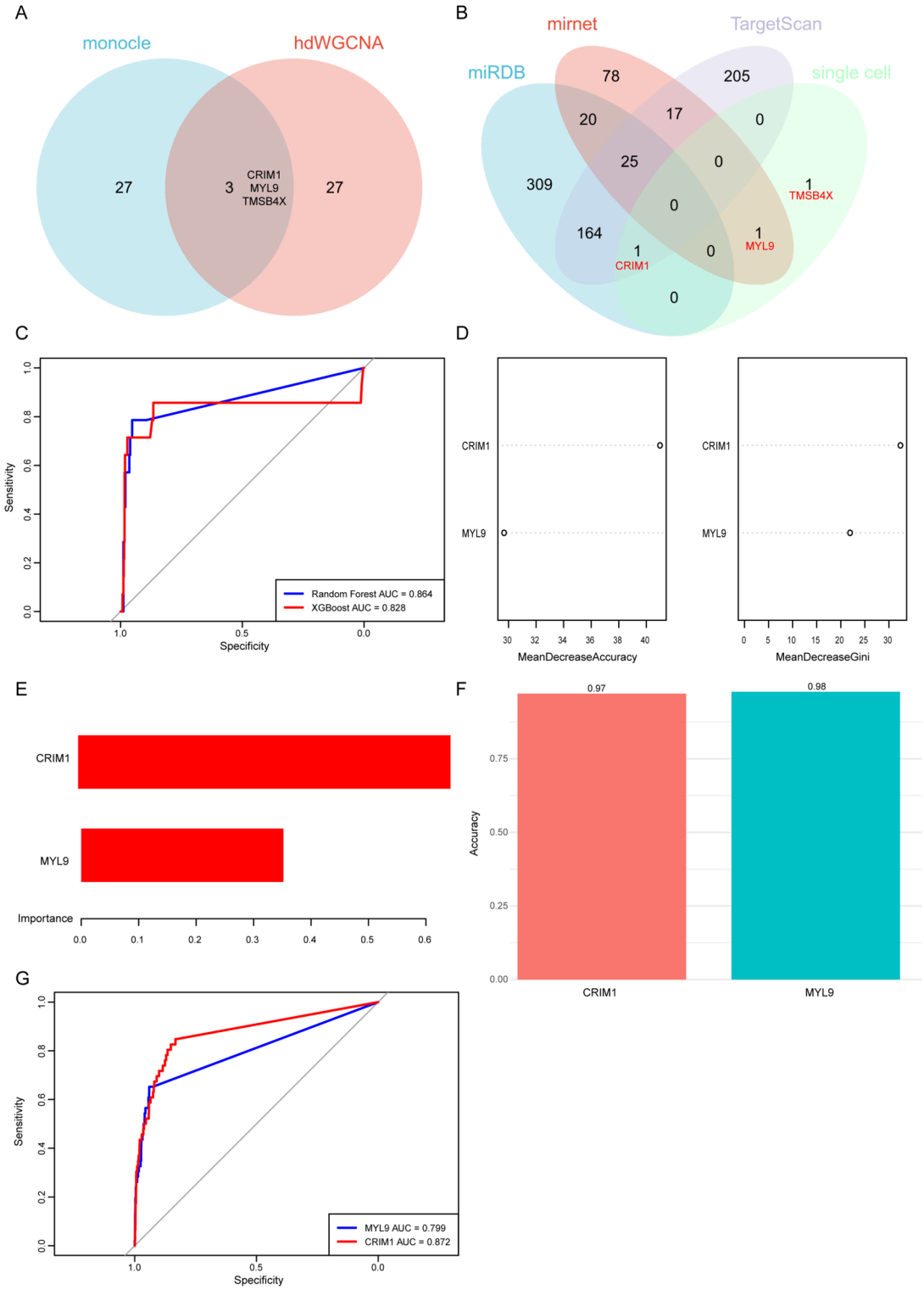

Fig. 6. Identification of miR-223-3p target genes in HBV-MN through integrated bioinformatics and machine learning. (**A**) Venn diagram comparing genes identified via pseudotime (Monocle) and hdWGCNA analyses, highlighting overlapping candidates (CRIM1, MYL9, TMSB4X). (**B**) Refinement of miR-223-3p targets using three databases (miRDB, miRnet, TargetScan) plus single-cell data, converging on CRIM1 and MYL9 as final candidate targets. (**C**) Evaluation of machine learning models (XGBoost, RF) performance via ROC curves. (**D**) Impact on HBV-MN predictive accuracy and gini impurity in RF model. (**E**) Comparative Gene Importance in HBV-MN with XGboost. (**F**) KNN algorithm validated the accuracy of two genes as predictive biomarkers in HBV-MN. (**G**) ROC analysis showed CRIM1 as the best predictive biomarker for HBV-MN in validation set. *RF* random forest; *ROC* receiver operating characteristic; *KNN* k-nearest neighbors.

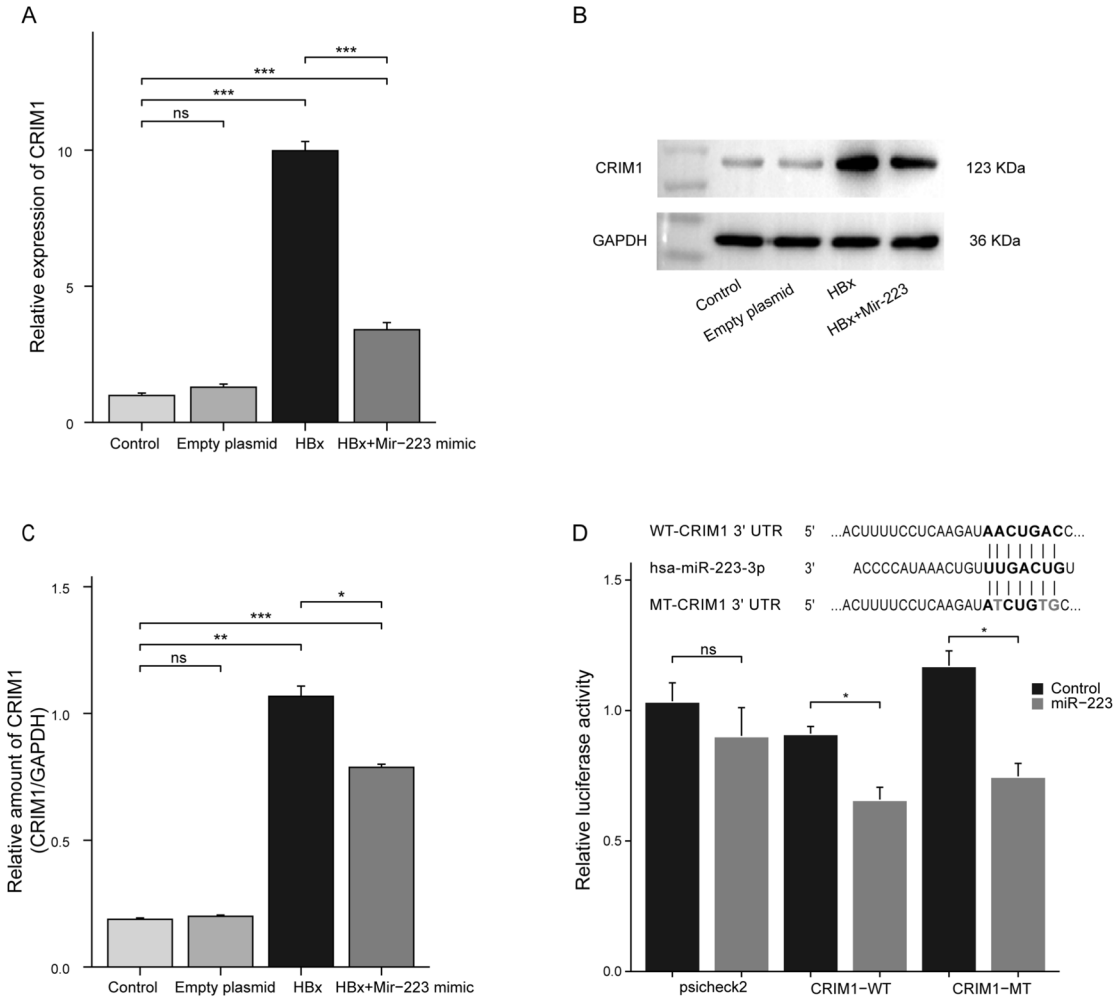

Fig. 7. The miR-223-3p overexpression downregulates CRIM1 in HBV-MN podocytes. (**A**) qPCR analysis of CRIM1 mRNA levels in control podocytes (Control, Empty plasmid) compared to those transfected with HBx alone or co-transfected with HBx and miR-223-3p mimic. (**B**) Western blot analysis demonstrating the regulation of CRIM1 protein levels by miR-223-3p. (**C**) Densitometric quantification of CRIM1 protein bands from the Western blot, confirming significant reduction in CRIM1 levels upon miR-223-3p overexpression. (**D**) Dual-luciferase reporter assay results illustrating fluorescence intensity in wild-type and mutant CRIM1, indicating miR-223-3p binding efficiency and suggesting complex interactions with the target gene. (ns: no significance, *p < 0.05, **p < 0.01, ***p < 0.001.).

on CRIM1 (Fig. 7B). Quantification of western blot bands showed a clear decrease in CRIM1 protein levels with miR-223-3p overexpression (Fig. 7C). We predicted two potential miR-223-3p binding sites on CRIM1: Position 193–199 of CRIM1 3'UTR (highly conserved) and Position 1628–1634 of CRIM1 3'UTR (poorly conserved). We chose the conserved site for our luciferase assay. Further investigation using a dual-luciferase reporter assay revealed a decrease in fluorescence intensity in wild-type CRIM1, indicating successful miR-223-3p targeting. Interestingly, the mutant CRIM1 group also showed reduced fluorescence, suggesting that the second potential binding site may contribute to the observed effects, indicating the possibility of alternative binding sites or indirect regulatory mechanisms involving miR-223-3p (Fig. 7D). These findings emphasize the regulatory role of miR-223-3p in CRIM1 expression and its potential impact on HBV-MN pathology. Original western blot images are presented in Supplementary files.

Discussion

In this study, we delve into the protective role of miR-223-3p and the potential role of CRIM1 as its target gene in HBV-MN. Utilizing single-cell RNA sequencing and advanced machine learning algorithms, we not only revealed the critical function of miR-223-3p in the progression of HBV-MN but also identified CRIM1 as a significant target, offering new insights into the pathophysiology of HBV-MN.

MiR-223-3p, a highly conserved microRNA expressed across various cell types including myeloid cells, hepatocytes, and cardiomyocytes, was initially thought to primarily play a role in hematopoiesis, immune responses, and inflammatory diseases. Subsequent research uncovered its significant role in the development of multiple diseases, including cardiovascular diseases where miR-223-3p is associated with atherosclerosis, vascular

remodeling, abnormal platelet reactivity, and myocardial ischemia[26,27]. In oncology, miR-223-3p's expression pattern varies across different cancers, acting as a tumor suppressor by targeting oncogenes like E2 F8 and ANLN, but also exhibiting oncogenic functions, as seen in prostate cancer through complex interactions with its strand partner, miR-223-5p[28]. In autoimmune diseases such as rheumatoid arthritis (RA), miR-223-3p negatively regulates granulocyte differentiation and activation and reduces E2 F1 protein expression, thereby affecting T lymphocyte function and participating in inflammation regulation[29].

In the field of kidney disease research, miR-223-3p plays a multifaceted and complex role, involving the regulation of immune responses, promotion or inhibition of inflammatory processes, and influencing kidney injury and repair mechanisms. Acute kidney injury studies have revealed dysregulated miR-223-3p as a contributor to immune cell infiltration and tissue remodeling[30]. Previous studies have shown that miR-223-3p plays a protective role in AKI induced by sepsis through interaction with NLRP3[31]. Long non-coding RNA DLX6-AS1, by sequestering miR-223-3p and promoting NLRP3 expression, exacerbated sepsis-induced tubular epithelial cell pyroptosis, whereas overexpression of miR-223-3p could partially reverse this process[32]. Similarly, in kidney stone disease, miR-223-3p attenuated calcium oxalate (CaOx)-induced tubular epithelial cell injury and reduced inflammatory responses and oxidative stress, suggesting its protective role in preventing kidney stone formation and progression[33]. The downregulation of miR-223-3p in CKD and DKD correlates with endothelial cell dysfunction and exacerbation of disease progression[34].

By targeting various molecular pathways, such as IL6ST/STAT3, overexpression of miR-223-3p could mitigate endothelial cell damage in a high-glucose environment, indicating its potential role in maintaining renal microenvironment stability and preventing CKD progression[6]. In diabetic kidney disease (DKD), high glucose environments also suppress miR-223-3p, leading to elevated intracellular calcium levels and ferroptosis via ITPR3 upregulation[35]. Furthermore, miR-223-3p has been found to regulate apoptosis and autophagy processes in renal cells, participating in renal protection mechanisms. For instance, in renal ischemia/reperfusion injury models, exosomes derived from bone marrow stem cells (BMSCs) delivering miR-223-3p inhibited NLRP3 expression, reduced inflammation and cell pyroptosis, and promoted the repair of damaged kidneys[36]. Lastly, miR-223-3p is closely related to inflammatory responses and immune regulation processes in disease, inhibiting the production of inflammatory cytokines and regulating the function of immune cells, thus playing a protective role in nephritis and autoimmune kidney diseases[37,38]. Our study further expands the knowledge on the function of miR-223-3p, highlighting its protective role in HBV-MN and emphasizing its potential in regulating CRIM1 expression. Nonetheless, precise regulation of miR-223-3p expression and the translation of these research findings into clinical applications require further investigation and exploration.

CRIM1 (Cysteine-rich motor neuron 1 protein), a cysteine-rich transmembrane protein expressed on the cell membrane through N-glycosylation modification, contains multiple von Willebrand factor type C domains, cysteine-rich repeats (CRRs), and an IGF-binding motif, indicating direct interactions with various signaling molecules and growth factors such as BMPs and members of the TGF-β superfamily[39]. These interactions regulate cell proliferation, migration, and differentiation, making CRIM1 crucial in embryonic development, organ formation, and disease progression[40]. Previous studies have shown that viral infections can lead to changes in CRIM1 levels[41]. Elevated CRIM1 expression has been observed in cardiomyocytes of individuals with coronary heart disease and in bronchial epithelial cells of COPD patients[42,43]. In cancer, CRIM1 influences tumor cell invasiveness and migration by regulating the composition of the extracellular matrix and cell–cell interactions. For example, CRIM1 enhances renal cancer cell migration and invasion by reducing E-cadherin expression and increasing Claudin-1 and matrix metalloproteinases (MMPs)[44]. Additionally, CRIM1 has been found to be a target of the Hippo pathway, affecting adhesion and migration of lung and gastric cancer cells, showcasing its multifunctionality in regulating cancer progression[45,46].

Genome-wide studies have suggested CRIM1 as a locus associated with kidney disease biomarkers[47]. In kidney-related diseases, CRIM1's role has been proven to be multifaceted, including maintaining the structure and function of podocytes during kidney development and preserving glomerular integrity in adult kidney disease models[48,49]. The interaction between CRIM1 and BMPs is particularly important as it regulates the intensity of BMP signaling in kidney development, affecting kidney morphogenesis and cell differentiation[40]. For example, reduced CRIM1 expression is associated with podocyte injury, foot process effacement, and proteinuria, as CRIM1 is essential for maintaining normal signaling between podocytes and glomerular endothelial cells[39]. Additionally, CRIM1's role in maintaining angiogenesis and microvascular integrity further emphasizes its central position in kidney health and disease[50]. Our findings linking CRIM1 to HBV-MN provide new insights into its role in kidney disease and may offer new targets for disease treatment. Notably, a recent integrated proteomic and metabolomic study identified a module—encompassing podocyte-expressed CRIM1—that was strongly associated with kidney failure in chronic kidney disease. This study highlighted that transmembrane–ephrin receptor activity, a pathway in which CRIM1 is involved, had one of the highest odds ratios for kidney disease progression. These findings lend further support to our data, suggesting that CRIM1 not only plays a pivotal role in maintaining podocyte integrity but also may serve as a potential therapeutic target in the context of HBV-MN and broader renal diseases[51].

despite the significant insights provided by our study regarding the roles of miR-223-3p and CRIM1 in HBV-MN, several limitations exist. First, the limited sample size may affect the general applicability of our results. Although our methods, including single-cell sequencing analysis and machine learning algorithms, showcased high precision and depth, future studies should include more biological replicates and diverse patient samples to further validate and expand these findings. Second, while we identified CRIM1 as a potential target of miR-223-3p and suggested its potential in HBV-MN treatment, causal relationships cannot be directly inferred. One avian model has demonstrated that miR-223 can indeed target CRIM1, suggesting an evolutionarily conserved regulatory axis. This observation aligns with our current findings, where overexpression of miR-223-3p in an HBV-MN model leads to reduced CRIM1 expression, implicating CRIM1 as a potentially critical downstream

effector of miR-223-3p in renal disease contexts[52]. Experimental validation, such as knockout and overexpression experiments, would be valuable to elucidate the exact mechanisms of these molecules in HBV-MN. Moreover, these findings need to be verified through further in vitro and in vivo experiments. Additionally, considering the role of miR-223-3p in various biological processes and diseases, further exploration of other signaling pathways and mechanisms involving miR-223-3p and CRIM1, as well as their interactions with other potential target genes, will be an important direction for future research. Furthermore, the miR-223-3p guide strand forms a microRNA-induced silencing complex to execute target gene silencing, while the lagging strand from the double-stranded miRNA is degraded without controlling target gene expression. miR-223-5p is the lagging strand of the miR-223 duplex[28]. Recent findings suggest that miR-223-3p and miR-223-5p target different molecules and pathways in the same disease context, indicating complementary or opposing actions. For instance, in myocardial ischemia/reperfusion injury, miR-223-3p and miR-223-5p co-regulate necroptotic and inflammatory pathways by targeting different mediators (e.g., TNFR1, DR6 (miR-223-5p), NLRP3, and IκB kinase α (miR-223-3p))[53]. Additionally, while miR-223-3p displayed oncogenic functions in PCa, subsequent research characterized the tumor-suppressive capabilities of miR-223-5p[28]. MiR-223-3p and miR-223-5p exhibit a nuanced relationship, suggesting that they can act independently and cooperatively under various conditions. How the miR-223 duplex influences HBV-MN disease progression and physiological responses requires further research.

Lastly, the role and mechanisms of CRIM1 deserve in-depth study to clarify its specific function in the progression of HBV-MN. Notably, circCRIM1, a circular RNA produced by the CRIM1 gene through back-splicing, plays a significant role in regulating gene expression, particularly as a microRNA sponge, affecting the availability of miRNAs and the expression of their target genes[54]. CircCRIM1, by binding specific miRNAs, regulates the expression of multiple target genes, including CRIM1, thereby playing roles in various biological processes, including cell proliferation, differentiation, and disease progression regulation. CircCRIM1, by sequestering miR-146a-5p and miR-342-3p, upregulates the expression of NUMB and TCF12, promoting cancer cell epithelial-mesenchymal transition (EMT) and tumor progression, with notable effects in esophageal squamous cell carcinoma, ovarian cancer, and nasopharyngeal carcinoma[54–56]. Research on the role of circCRIM1 in kidney disease is relatively scarce, but based on the functions of circRNAs in other tissues and diseases, it can be inferred that circCRIM1 might play a role in kidney disease through similar mechanisms. For example, by affecting miRNA activity, it might influence kidney cell injury response, cell survival and apoptosis, and inflammation. Future research should focus on the potential roles and mechanisms of CRIM1 and circCRIM1 in kidney disease, especially in renal dysplasia and chronic kidney disease.

Conclusion

In summary, our study emphasizes the important roles of miR-223-3p and CRIM1 in HBV-MN, providing new insights that contribute to the understanding of the disease's molecular mechanisms and the development of potential therapeutic strategies. By combining high-throughput techniques and advanced data analysis methods, our research not only enhances our understanding of the pathophysiology of HBV-MN but also offers new research tools and strategies for the field of kidney disease research.

Data availability

Publicly available datasets were analyzed in this study. This data can be found here: https://www.ncbi.nlm.nih.gov/geo/. Data supporting this study are available in the GEO repository under accession numbers [GSE199850, GSE171458]. All results of the experiments and bioinformatics analysis were uploaded to Supplementary Material.

Received: 19 July 2024; Accepted: 21 May 2025
Published online: 29 May 2025

References

1. Fatehi, F., Bingham, R. J., Stockley, P. G. & Twarock, R. An age-structured model of hepatitis B viral infection highlights the potential of different therapeutic strategies. *Sci. Rep.* **12**, 1252 (2022).
2. Lan, W. et al. Metabolic regulation of hepatitis B virus infection in HBV-transgenic mice. *Metabolites* **12**, 287 (2022).
3. Lei, X. et al. Hepatitis B virus X protein decreases nephrin expression and induces podocyte apoptosis via activating STAT3. *Exp. Ther. Med.* **17**, 7453 (2019).
4. Saravanan, S. et al. miRNA-24 and miRNA-466i-5p controls inflammation in rat hepatocytes. *Cell Mol. Immunol.* **12**, 113–115 (2015).
5. Feng, M. et al. HBx-induced PLA2R overexpression mediates podocyte pyroptosis through the ROS-NLRP3 signaling pathway. *Renal Fail.* **45**, 808 (2023).
6. Tang, P. et al. miR-223-3p mediates the diabetic kidney disease progression by targeting IL6ST/STAT3 pathway. *Biochem. Biophys. Res. Commun.* **648**, 50–58 (2023).
7. Fava, A. et al. Urine proteomics and renal single-cell transcriptomics implicate interleukin-16 in lupus nephritis. *Arthritis Rheumatol.* **74**, 829–839 (2022).
8. Wu, H. et al. Mapping the single-cell transcriptomic response of murine diabetic kidney disease to therapies. *Cell Metab.* **34**, 1064-1078.e6 (2022).
9. Yu, L. et al. Intrarenal single-cell sequencing of hepatitis B virus associated membranous nephropathy. *Front. Med.* **9**, 284 (2022).
10. Xu, J. et al. Single-cell profiling reveals transcriptional signatures and cell-cell crosstalk in anti-PLA2R positive idiopathic membranous nephropathy patients. *Front. Immunol.* **12**, 330 (2021).
11. Stuart, T. et al. Comprehensive integration of single-cell data. *Cell* **177**, 1888-1902.e21 (2019).
12. Aran, D. et al. Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nat. Immunol.* **20**, 163–172 (2019).
13. Morabito, S. et al. Single-nucleus chromatin accessibility and transcriptomic characterization of Alzheimer's disease. *Nat. Genet.* **53**, 1143–1155 (2021).

14. Morabito, S., Reese, F., Rahimzadeh, N., Miyoshi, E. & Swarup, V. hdWGCNA identifies co-expression networks in high-dimensional transcriptomics data. *Cell Rep. Methods* **3**, 100498 (2023).
15. Qiu, X. et al. Reversed graph embedding resolves complex single-cell trajectories. *Nat. Methods* **14**, 979–982 (2017).
16. Jin, S. et al. Inference and analysis of cell-cell communication using Cell Chat. *Nat. Commun.* **12**, 1088 (2021).
17. Chen, Y. & Wang, X. miRDB: an online database for prediction of functional microRNA targets. *Nucleic Acids Res.* **48**, D127–D131 (2020).
18. Fan, Y. et al. miRNet - dissecting miRNA-target interactions and functional associations through network-based visual analysis. *Nucleic Acids Res.* **44**, W135–W141 (2016).
19. Agarwal, V., Bell, G. W., Nam, J. W. & Bartel, D. P. Predicting effective microRNA target sites in mammalian mRNAs. *Elife* **4**, 5005 (2015).
20. Ishwaran, H. The effect of splitting on random forests. *Mach. Learn.* **99**, 75–118 (2015).
21. Deng, X., Li, M., Deng, S. & Wang, L. Hybrid gene selection approach using XGBoost and multi-objective genetic algorithm for cancer classification. *Med. Biol. Eng. Comput.* **60**, 663–681 (2022).
22. Houssein, E. H. & Sayed, A. A modified weighted mean of vectors optimizer for Chronic Kidney disease classification. *Comput. Biol. Med.* **155**, 106691 (2023).
23. Kanehisa, M., Furumichi, M., Sato, Y., Matsuura, Y. & Ishiguro-Watanabe, M. KEGG: biological systems database as a model of the real world. *Nucleic Acids Res.* **53**, D672–D677 (2025).
24. Kanehisa, M. Toward understanding the origin and evolution of cellular organisms. *Protein Sci.* **28**, 1947–1951 (2019).
25. Kanehisa, M. & Goto, S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).
26. Zhang, M. W., Shen, Y. J., Shi, J. & Yu, J. G. MiR-223–3p in cardiovascular diseases: A biomarker and potential therapeutic target. *Front. Cardiovasc. Med.* **7**, 10561 (2021).
27. Costa, A. C. et al. Characterizing the inflammatory microenvironment in K14-HPV16 transgenic mice: mast cell infiltration and microRNA expression. *Cancers* **14**, 2216 (2022).
28. Wei, Y. et al. miR-223-5p targeting ERG inhibits prostate cancer cell proliferation and migration. *J Cancer.* **11**, 4453–4463 (2020).
29. Reyes-Long, S. et al. Nociceptive related microRNAs and their role in rheumatoid arthritis. *Mol. Biol. Rep.* **47**, 7265–7272 (2020).
30. Zhang, J. N. et al. Integrated analysis of gene expression and immune cell infiltration reveals dysregulated genes and miRNAs in acute kidney injury. *Mol. Biotechnol.* https://doi.org/10.1007/s12033-024-01344-x (2024).
31. Xie, Z. et al. Human bone marrow mesenchymal stem cell-derived extracellular vesicles reduce inflammation and pyroptosis in acute kidney injury via miR-223-3p/HDAC2/SNRK. *Inflamm. Res.* **72**, 553–576 (2023).
32. Tan, J., Fan, J., He, J., Zhao, L. & Tang, H. Knockdown of LncRNA DLX6-AS1 inhibits HK-2 cell pyroptosis via regulating miR-223-3p/NLRP3 pathway in lipopolysaccharide-induced acute kidney injury. *J. Bioenerg. Biomembr.* **52**, 367–376 (2020).
33. Lv, P. et al. XIST Inhibition Attenuates Calcium Oxalate Nephrocalcinosis-Induced Renal Inflammation and Oxidative Injury via the miR-223/NLRP3 Pathway. *Oxid Med Cell Longev.* **2021**, 1676152 (2021).
34. Ulbing, M. et al. MicroRNAs 223–3p and 93–5p in patients with chronic kidney disease before and after renal transplantation. *Bone* **95**, 115–123 (2017).
35. Wang, D. et al. High glucose elevates intracellular calcium level and induces ferroptosis in glomerular endothelial cells through the miR-223–3p/ITPR3 pathway. *Mol. Cell Endocrinol.* **594**, 112384 (2024).
36. Sun, Z. et al. MSC-derived extracellular vesicles activate mitophagy to alleviate renal ischemia/reperfusion injury via the miR-223–3p/NLRP3 Axis. *Stem Cells Int.* **2022**, 1–20 (2022).
37. Tang, H. C., Lai, Y. Y., Zheng, J., Jiang, H. Y. & Xu, G. miR-223–3p inhibits antigen endocytosis and presentation and promotes the tolerogenic potential of dendritic cells through targeting mannose receptor signaling and rhob. *J. Immunol. Res.* **2020**, 1–17 (2020).
38. Loukachov, V. V. et al. Identification of liver and plasma microRNAs in chronic hepatitis B virus infection. *Front. Cell Infect. Microbiol.* **12**, 964 (2022).
39. Nyström, J. et al. CRIM1 is localized to the podocyte filtration slit diaphragm of the adult human kidney. *Nephrol. Dial. Transplant.* **24**, 2038–2044 (2009).
40. Zeng, H. & Tang, L. CRIM1, the antagonist of BMPs, is a potential risk factor of cancer. *Curr. Cancer Drug Targets* **14**, 652–658 (2014).
41. Mo, X. et al. Microarray profiling analysis uncovers common molecular mechanisms of rubella virus, human cytomegalovirus, and herpes simplex virus type 2 infections in ECV304 cells. *Curr. Mol. Med.* **11**, 481–488 (2011).
42. Hill, M. C. et al. Integrated multi-omic characterization of congenital heart disease. *Nature* **608**, 181–191 (2022).
43. Di Stefano, A. et al. Bone morphogenic proteins and their antagonists in the lower airways of stable COPD patients. *Biology* **12**, 1304 (2023).
44. Ogasawara, N. et al. Reduction of membrane protein CRIM1 decreases E-cadherin and increases claudin-1 and MMPs, enhancing the migration and invasion of renal carcinoma cells. *Biol. Pharm. Bull.* **41**, 604–611 (2018).
45. Zeng, H. et al. CRIM1, a newfound cancer-related player, regulates the adhesion and migration of lung cancer cells. *Growth Factors* **33**, 384–392 (2015).
46. Lim, B. et al. Integrative genomics analysis reveals the multilevel dysregulation and oncogenic characteristics of TEAD4 in gastric cancer. *Carcinogenesis* **35**, 1020–1027 (2014).
47. Thomson, R. J. et al. New genetic loci associated with chronic kidney disease in an indigenous Australian population. *Front. Genet.* **10**, 330 (2019).
48. Wilkinson, L. et al. Crim1KST264/KST264 mice implicate Crim1 in the regulation of vascular endothelial growth factor-A activity during glomerular vascular development. *J. Am. Soc. Nephrol.* **18**, 1697–1708 (2007).
49. Wilkinson, L. et al. Loss of renal microvascular integrity in postnatal Crim1 hypomorphic transgenic mice. *Kidney Int.* **76**, 1161–1171 (2009).
50. Nakashima, Y. & Takahashi, S. Induction of cysteine-rich motor neuron 1 mRNA expression in vascular endothelial cells. *Biochem. Biophys. Res. Commun.* **451**, 235–238 (2014).
51. Schlosser, P. et al. Association of integrated proteomic and metabolomic modules with risk of kidney disease progression. *J. Am. Soc. Nephrol.* **35**, 923–935 (2024).
52. Xiang, J. et al. MiR-223 inhibits proliferation and steroid hormone synthesis of ovarian granulosa cell via the AKT signaling pathway by targeting CRIM1 in chicken. *Poult. Sci.* **103**, 103910 (2024).
53. Qin, D. et al. MicroRNA-223-5p and -3p cooperatively suppress necroptosis in ischemic/reperfused hearts. *J. Biol. Chem.* **291**, 20247–20259 (2016).
54. Du, Y. et al. CircCRIM1 promotes ovarian cancer progression by working as ceRNAs of CRIM1 and targeting miR-383–5p/ZEB2 axis. *Reprod. Biol. Endocrinol.* **19**, 176 (2021).
55. Li, X. P. et al. Circular RNA hsa_circ_0002938 (circCRIM1) promotes the progression of esophageal squamous cell carcinoma by upregulating transcription factor 12. *Neoplasma* **70**, 145–157 (2023).
56. Hong, X. et al. Circular RNA CRIM1 functions as a ceRNA to promote nasopharyngeal carcinoma metastasis and docetaxel chemoresistance through upregulating FOXQ1. *Mol. Cancer* **19**, 33 (2020).

Acknowledgements
The authors would like to thank the staff at the Affiliated Hospital of Qingdao University Biological Laboratory for their assistance in separating and submitting test samples.

Author contributions
YZH, QA, XXY, and WJ conceived and designed the study. YZH carried out all experiments and wrote the manuscript. YZH, QA and WJ participated in bioinformatics analysis. WJ, XXY supported the study. QA and WJ supervised the study. YZH, QA, XXY, and WJ revised the manuscript. All authors read and approved the final manuscript.

Funding
This research was supported by grants from the National Natural Science Foundation of China [82370724], the Qingdao Key Health Discipline Development Fund and the Qingdao Key Clinical Specialty Elite Discipline project.

Competing interests
The authors declare no competing interests.

Ethical approval
In conducting this study, no human participants or samples were directly involved in our research. We utilized publicly accessible GEO datasets, ethical clearance for the original data collection was duly obtained by each contributing study's institutional review board, where all participants provided informed consent. This study has been registered and approved by the ethics review board of ethics committee approval of the affiliated hospital of Qingdao university. Ethics approval number: QYFY WZLL 28,663.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-03625-0.

Correspondence and requests for materials should be addressed to W.J.