EPIGENOMICS
2025, VOL. 17, NO. 14, 935–952
https://doi.org/10.1080/17501911.2025.2548762

Taylor & Francis
Taylor & Francis Group

RESEARCH ARTICLE

OPEN ACCESS Check for updates

The NEAT1/miR-124-3p/CCL2 axis in chronic kidney disease progression: integrated bioinformatics analysis and experimental validation

Guanting Chen[a,b], Linqi Zhang[a,b], Yaoxian Wang[b], Jianfeng Wang[b], Kang Yang[a,b], Xixi Wang[a] and Xu Chen[a]

[a]Department of Nephrology, First Affiliated Hospital of Henan University of Chinese Medicine, Zhengzhou, Henan Province, China; [b]Henan Collaborative Innovation Center for Prevention and Treatment of Major Diseases through Integrated Chinese and Western Medicine, Henan University of Chinese Medicine, Zhengzhou, Henan Province, China

ABSTRACT

Background: Chronickidney disease (CKD) is a major global health burden lacking effectivetherapies. Renal interstitial fibrosis (RIF) is a key pathological driver ofCKD progression. This study aimed to identify novel diagnostic biomarkers and therapeutictargets.

Research design and methods: Weanalyzed the GEO dataset GSE137570 to identify differentially expressed genes(DEGs). Protein-protein interaction (PPI) networks were constructed to screen HubGenes. A competing endogenous RNA (ceRNA) network was predicted. Validationincluded single-cell sequencing, in vitro epithelial-mesenchymal transition(EMT) models using Transforming growth factor-β 1 (TGF-β1)-treated TCMK1 cells, clinical samples (64 CKD patients, 20 healthy controls), and dual-luciferasereporter assays (DLRA).

Results: FiveHub Genes (EGF, VCAN, CXCL1, MMP7, CCL2) were identified, with CCL2 being themost central. Enrichment analyses linked them to immune/inflammatory responses.DLRA confirmed specific targeting between miR-124-3p and both NEAT1 and CCL2,supporting the NEAT1/miR-124-3p/CCL2 axis. Clinically, serum CCL2 increasedwhile miR-124-3p and NEAT1 decreased with CKD progression; all three showedgood diagnostic accuracy for staging.

Conclusions: EGF,VCAN, CXCL1, MMP7, and particularly CCL2 are potential CKDbiomarkers/therapeutic targets. The NEAT1/miR-124-3p/CCL2 axis is a keyregulatory pathway in CKD. Key limitations include the moderate sample sizes inbioinformatics and clinical cohorts.

PLAIN LANGUAGE SUMMARY

Doctors lack good tests to track how chronic kidneydisease (CKD) gets worse over time. By studying kidney patient data, we firstfound five important substances in the body linked to CKD getting worse. One ofthese, called CCL2, had the strongest connection to the disease.

To learn how CCL2 is controlled, we studied two othermolecules that might affect it: NEAT1 (a longer molecule) and miR-124-3p (a smaller one). Our tests showed that NEAT1 "catches" miR-124-3p, stopping itfrom reducing CCL2. When we checked blood levels in 64 CKD patients, we sawthat as kidney function got worse, NEAT1 and miR-124-3p went down while CCL2went up.

These findings could lead to new blood tests tomonitor CKD and may help develop future treatments.

ARTICLE HISTORY

Received 25 April 2025
Accepted 11 August 2025

KEYWORDS

Bioinformatics; renal interstitial fibrosis; chronic kidney disease; NEAT1; miR-124-3p; CCL2

1. Introduction

Chronic kidney disease (CKD) represents a significant global public health challenge. Epidemiological studies indicate that CKD accounts for approximately 1.2 million annual deaths worldwide, with mortality rates continuing to rise. Projections suggest it will become the fifth leading cause of death by 2040 [1,2].

Renal interstitial fibrosis (RIF) constitutes both the common pathological pathway and primary pathological basis for progression of CKD to end-stage renal disease (ESRD) [3,4]. Accumulating evidence demonstrates that the severity of tubulointerstitial injury serves as a critical determinant of renal functional decline and prognosis. Notably, CKD progression and outcomes correlate more strongly with the presence and extent of tubulointerstitial lesions than with glomerular damage [5]. Tubulointerstitial injury is a pivotal predictor of 10-year renal survival, exerting a substantially greater pathological influence than glomerular lesions [6,7].

Renal tubular epithelial cells (TECs), as primary targets of metabolic, immune, ischemic, and toxic insults, mount diverse responses to cellular stress or injury. These include proliferation, autophagy, growth arrest, epithelial-mesenchymal transition (EMT), and various forms of cell death. Such responses trigger inflammatory cell infiltration, fibroblast activation/proliferation, extracellular matrix (ECM) deposition, tubular atrophy, and microvascular rarefaction, thereby initiating RIF progression. This cascade ultimately drives decline in glomerular filtration rate (GFR) and progression to ESRD [8,9].

CONTACT Yaoxian Wang wyx3203@sina.com Henan Collaborative Innovation Center for Prevention and Treatment of Major Diseases through Integrated Chinese and Western Medicine, Henan University of Chinese Medicine, Zhengzhou, Henan Province 450003, China; Linqi Zhang 1478404656@qq.com Department of Nephrology, First Affiliated Hospital of Henan University of Chinese Medicine, Zhengzhou, Henan Province 450003, China

Supplemental data for this article can be accessed online at https://doi.org/10.1080/17501911.2025.2548762

Article highlights

- Identified EGF, VCAN, CXCL1, MMP7, and CCL2 as key Hub Genes and potential diagnostic biomarkers/therapeutic targets for CKD progression.
- Revealed NEAT1/miR-124-3p/CCL2 as a novel and functionally validated ceRNA regulatory axis driving renal inflammation and fibrosis in CKD.
- Demonstrated significant alterations in renal immune cell infiltration patterns across CKD stages and their correlation with Hub Gene expression.
- Validated the dysregulation of serum CCL2, miR-124-3p, and NEAT1 levels in CKD patients, showing strong correlation with renal function decline (e.g., Scr, eGFR).
- Established the diagnostic potential of serum CCL2, miR-124-3p, and NEAT1, particularly for distinguishing CKD progression stages, using ROC curve analysis.
- Proposed targeting the NEAT1/miR-124-3p/CCL2 axis (e.g., via NEAT1 inhibition) as a promising therapeutic strategy for CKD.

Consequently, elucidating the molecular mechanisms underlying RIF pathogenesis and developing effective therapeutic interventions are crucial for delaying CKD progression and reducing ESRD burden. However, CKD presents dual clinical challenges due to its insidious onset, multifactorial etiology, and heterogeneous pathogenesis: limited early diagnostic capability and inadequate therapeutic options. Current clinical guidelines advocate multi-target therapies including angiotensin-converting enzyme inhibitors (ACEIs), angiotensin receptor blockers (ARBs), sodium-glucose cotransporter-2 inhibitors (SGLT2is), and mineralocorticoid receptor antagonists (MRAs) [10]. Despite these approaches, existing interventions inadequately suppress the core pathological processes driving RIF [11]. Diagnostically, invasive renal biopsy carries risks of sampling variability and observer bias, while conventional markers – such as ultrasound imaging, blood urea nitrogen (BUN), and serum creatinine (Scr) – lack sufficient sensitivity and are susceptible to confounding factors. These limitations impede dynamic monitoring of early renal injury [12]. Thus, there is an urgent need to develop noninvasive biomarkers with high sensitivity and specificity to overcome current diagnostic constraints.

Emerging evidence indicates that long non-coding RNAs (lncRNAs) and microRNAs (miRNAs) are detectable in diverse biofluids (e.g., serum, plasma, urine). Their expression profiles may reflect systemic responses to pathophysiological stress, mediating cellular activation, proliferation, transformation, apoptosis, lipid metabolism, and renal fibrosis progression [13]. As early-response molecules to renal injury, they possess unique clinical utility and represent promising supplementary tools for broader CKD detection [14,15]. Mechanistically, lncRNAs can function as competing endogenous RNAs (ceRNAs) by sequestering miRNAs via molecular sponge activity, thereby inhibiting miRNA-mediated regulation of target genes [16]. Among these, nuclear enriched abundant transcript 1 (NEAT1), a recently characterized lncRNA, serves as an essential architectural component of paraspeckles – subnuclear membraneless organelles [17]. Cumulative studies demonstrate that NEAT1 is markedly upregulated in fibrotic pathologies across multiple tissues, and its genetic ablation significantly attenuates fibrotic progression [18]. In renal disease contexts, NEAT1 activates signaling pathways including Akt/mTOR and ERK1/2, thereby regulating EMT and ECM protein expression [19,20]. It further modulates target genes by adsorbing specific miRNAs (e.g., miR-23c, miR-129, miR-31), consequently influencing inflammatory responses and renal fibrosis [21–23].

miR-124-3p, a critical miRNA expressed in immune cells and multiple organs, is activated by inflammatory signals and functions as a homeostatic negative regulator. Its dysregulation is recognized as a key modulator of immune cell function, inflammation, and immune responses [24]. Furthermore, miR-124-3p contributes to renal and fibrotic pathologies by targeting genes such as epithelial cadherin (E-cadherin) and vimentin, thereby regulating EMT in TECs [25]. This regulatory function aligns with core histopathological features of CKD: persistent renal inflammation and excessive ECM deposition [26].

Chemokine (C-C motif) ligand 2 (CCL2), a CC-subfamily cytokine secreted by diverse cell types (e.g., macrophages, monocytes, endothelial cells, fibroblasts, mesangial cells, TECs, smooth muscle cells), binds to its cognate receptor CCR2 to promote immune cell migration/activation, thereby exacerbating inflammation and fibrosis [27]. During renal injury, CCL2 released from TECs recruits bone marrow-derived monocytes/progenitors into circulation and chemoattracts CCR2+ monocytes to lesion sites. This process triggers profibrotic cytokine secretion and pathological tissue remodeling, ultimately inducing ECM accumulation in tubules and glomeruli, which drives RIF and glomerulosclerosis [28].

The Gene Expression Omnibus (GEO) database, curated by the National Center for Biotechnology Information (NCBI), represents a large-scale public repository archiving high-throughput gene expression data from global research institutions. It encompasses single- and dual-channel microarray experiments alongside non-array-based methodologies, including Serial Analysis of Gene Expression (SAGE), mass spectrometry proteomics, and high-throughput sequencing data [29]. In this study, we performed an integrated analysis of the GSE137570 dataset from GEO, which contains transcriptomic profiles of renal tissues from CKD patients across disease stages. Through screening for differentially expressed genes (DEGs) across progressive CKD stages, we identified CCL2 as a potential key pathogenic mediator and delineated its upstream regulatory axis involving NEAT1 and miR-124-3p. These bioinformatic findings were subsequently validated through clinical samples and in vitro experiments. This work aims to establish a molecular framework for evaluating and therapeutically intervening in CKD progression while providing novel insights for future research. The analytical workflow is summarized in Figure 1.

2. Materials and methods

2.1. Data sources and processing

Through the GEO database (https://www.ncbi.nlm.nih.gov/geo/) maintained by the NCBI, we retrieved the latest relevant data using the search terms "Chronic kidney disease," "Renal interstitial fibrosis," and "Homo sapiens," ultimately identifying the human renal tissue RNA sequencing dataset GSE137570—publicly submitted by Doyle R. and McEvoy C. (University

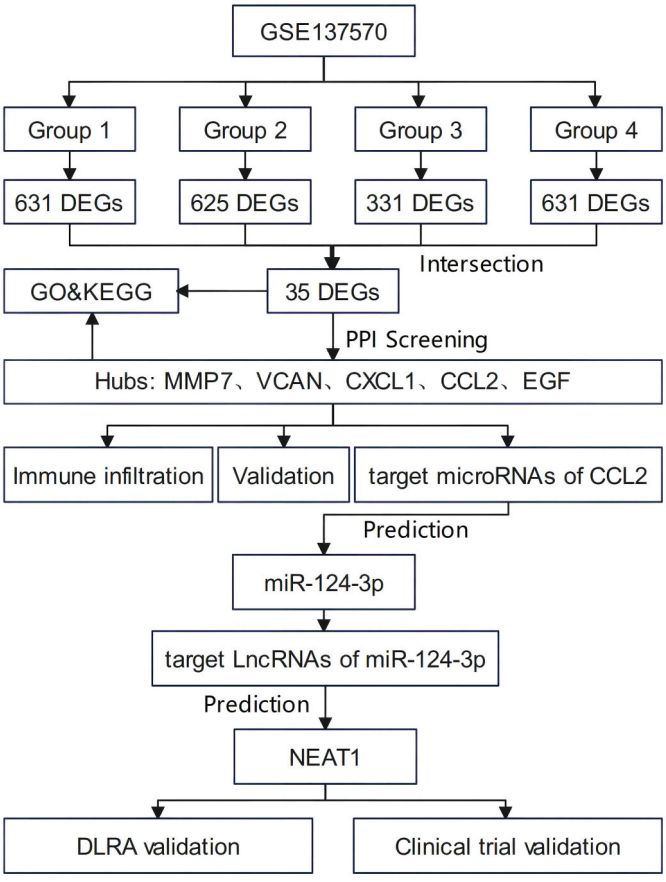

Figure 1. Schematic diagram of the integrated bioinformatics and experimental workflow for identifying the NEAT1/miR-124-3p/CCL2 axis in chronic kidney disease (CKD) progression. Key steps include: (1) data acquisition from gene expression omnibus (GEO) dataset GSE137570; (2) differential gene expression analysis; (3) enrichment and gene set variation analysis (GSVA) pathway analysis; (4) protein-protein interaction (PPI) network construction and Hub gene identification; (5) prediction of lncRNA-miRNA-mRNA regulatory network; (6) validation via single-cell sequencing (SCP975), in vitro EMT models, and clinical samples; (7) functional validation by dual-luciferase reporter assays.

College Dublin) on 24 October 2022. According to GEO's explicit data use policies, downloading and analyzing this publicly authorized dataset requires no additional permissions or written consent from data owners.

The dataset comprises two independent platforms: GPL23227 (BGISEQ-500) and GPL11154 (Illumina HiSeq 2500), which performed high-throughput sequencing on two separate cohorts of CKD patient renal tissue samples: Cohort 1 (n = 17) and Cohort 2 (n = 24). This dataset contains transcriptomic data from a total of 41 renal tissue samples obtained from CKD patients at various stages. Cohort 1 incorporated sex, age, and renal function progression status within a 60-month median follow-up period, with grouping based on whether patients exhibited renal function progression – reflecting a longitudinal analysis of long-term prognosis; Cohort 2 included sex, age, estimated glomerular filtration rate (eGFR), and tubulointerstitial fibrosis (TIF) scores. Patients were stratified into three independent comparison groups according to baseline eGFR levels: eGFR < 30 mL/min/1.73 m^2 versus ≥30 mL/min/1.73 m^2, eGFR < 60 mL/min/1.73 m^2 versus ≥60 mL/min/1.73 m^2, and eGFR < 90 mL/min/1.73 m^2 versus ≥90 mL/min/1.73 m^2. This stratification represents a cross-sectional analysis based on baseline renal function severity. Detailed group classifications and sample sizes are provided in Table 1.

The raw gene expression count matrix of dataset GSE137570 was downloaded from GEO as the foundational input data for subsequent analysis. Differential gene expression analysis across sample groups in GSE137570 was performed using NetworkAnalyst 3.0 (https://www.networkanalyst.ca/) with the following parameters: during data filtering and normalization, the Variance filter threshold was set to "15," the Low abundance threshold to "4," and the normalization method to "Log2-counts per million." For the statistical method, "DESeq2" was selected; in Study Design-Primary Factor, "CLASS" was chosen for grouping; in Specific comparison, the trial group was designated first followed by the control group. Screening criteria included a false discovery rate (FDR)-adjusted p value < 0.05 and |log2 Fold Change (FC)| > 1. The Benjamini-Hochberg FDR correction procedure was applied to control false positive risk arising from multiple hypothesis testing inherent in transcriptome-wide analyses. The intersection of DEGs filtered from each group was identified to explore targets highly associated with CKD progression.

2.2. Enrichment analysis

Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were performed on the intersected target genes using multiple databases: NetworkAnalyst (https://www.networkanalyst.ca/), Metascape (https://metascape.org), and WebGestalt (http://www.webgestalt.org/). GO enrichment analysis assessed Biological Process (BP), Molecular Function (MF), and Cellular Component (CC) categories. Terms with a significance level of p < 0.05 were considered statistically enriched.

2.3. Gene set variation analysis (GSVA) analysis

The MsigDB Hallmark gene sets for Groups 1–4 were subjected to GSVA analysis [30,31] using the Wei Shengxin database (https://www.bioinformatics.com.cn/). The GSVA enrichment scores were calculated for each condition, followed by differential analysis of enrichment scores between control and experimental groups, with subsequent heatmap visualization.

Table 1. Grouping of GSE137570 dataset.

Grouping	Cohort	Trial group	Control group
Group 1	Cohort 1 (n = 17)	Chronic kidney disease (CKD) progression (n = 9)	CKD non-progression (n = 8)
Group 2	Cohort 2 (n = 24)	Estimated glomerular filtration rate (eGFR) < 30 mL/min·1.73 m^2 (n = 9)	eGFR ≥30 mL/min·1.73 m^2 (n = 15)
Group 3	Cohort 2 (n = 24)	eGFR < 60 mL/min·1.73 m^2 (n = 16)	eGFR ≥60 mL/min·1.73 m^2 (n = 8)
Group 4	Cohort 2 (n = 24)	eGFR < 90 mL/min·1.73 m^2 (n = 21)	eGFR ≥90 mL/min·1.73 m^2 (n = 3)

2.4. Hub gene identification

To construct the Protein-Protein Interaction (PPI) network, the intersection set of DEGs across comparison groups was uploaded to the STRING database (version 12.0; https://cn.string-db.org/) using the default minimum interaction confidence score of 0.4. The resulting PPI network was visualized using Cytoscape software (version 3.9.1). Network topological parameters were calculated using Cytoscape's built-in Network Analyzer tool. Hub Genes within the PPI network were identified using two complementary Cytoscape plugins: (1) Molecular Complex Detection (MCODE): Identifies densely connected network regions using a vertex-weighting algorithm based on local node density. (2) CytoHubba: Ranks nodes by importance using 11 topological algorithms. Genes were prioritized using the Maximal Clique Centrality (MCC) method, which exhibits superior predictive accuracy. Hub Genes were defined as those consistently identified by both the MCODE and CytoHubba (MCC ranking) algorithms.

2.5. Construction of the lncRNA-miRNA-mRNA regulatory network

miRNAs targeting the identified Hub Genes were predicted using the MiRWalk (http://mirwalk.umm.uni-heidelberg.de/) and miRTarBase (https://miRTarBase.cuhk.edu.cn/) databases. The intersection of predicted miRNAs from both databases was retained for further analysis. Long non-coding RNAs (lncRNAs) interacting with these miRNAs were subsequently predicted using ENCORI (https://starbase.sysu.edu.cn/index.php) and MiRNET (https://www.mirnet.ca/). A comprehensive lncRNA-miRNA-mRNA interaction network was constructed and visualized using Cytoscape.

2.6. Immune infiltration analysis

Immune cell infiltration within renal tissues was characterized using the CIBERSORT algorithm, implemented via the "Immune Estimation" module of TIMER 2.0 (http://timer.comp-genomics.org/). This analysis was performed on two independent cohorts of CKD patient renal tissue samples from the GSE137570 dataset. The objectives were to: (1) quantify the relative abundance of distinct immune cell subtypes, and (2) investigate potential correlations between core target gene expression and differentially abundant immune cell populations. Resultant immune cell fraction estimates were exported to the Xiantao Academic platform (https://www.xiantao.love/) for downstream statistical analysis and visualization.

2.7. Single-cell RNA sequencing validation

The Single Cell Portal (https://singlecell.broadinstitute.org) functions as an integrated repository and analytical platform for single-cell sequencing data, enabling systematic interrogation of gene expression profiles, cellular heterogeneity, intercellular communication networks, and cell-state dynamics at single-cell resolution. To validate Hub Gene expression patterns, we queried the database using "Chronic kidney disease" and "Renal interstitial fibrosis" as keywords, identifying the single-cell RNA sequencing dataset SCP975 deposited by Chung et al. [32]—comprising 73,511 cells across eight annotated types (Endothelial, Arteriolar Endothelial, Immune, Mesangial, Myofibroblasts, Parietal Epithelial Cells, Podocytes, and Smooth Muscle Cells) and integrating five established murine kidney disease models: Adriamycin Nephropathy, Anti-Glomerular Basement Membrane Nephritis, CD2AP Knockout, and Diabetic Kidney Disease. This resource provides a comprehensive single-cell resolution framework for renal interstitial fibrosis (RIF) research; the dataset underwent original authors' standardized preprocessing including quality control, normalization, and batch correction with complete cell-type annotation, and per the Single Cell Portal Open Data Access Agreement, analysis requires no additional permissions. Using the platform's visualization tools, we conducted gene-wise screening through the Scatter feature (Explore module) to spatially map Hub Gene expression across cellular compartments, enabling comparison of expression variations between pathological states and cellular lineages.

2.8. Patient sources and specimen collection

Following the 2022 KDIGO Clinical Practice Guideline for Chronic Kidney Disease Management as the reference standard [33], we enrolled 64 CKD patients admitted to the Nephrology Department of the First Affiliated Hospital of Henan University of Chinese Medicine between February 2023 and July 2023, along with 20 healthy volunteers from concurrent physical examinations. Participants were categorized into the CKD group (G2: $n = 14$, G3: $n = 30$, G4: $n = 20$) and healthy control group. This study was approved by the Medical Ethics Committee of the First Affiliated Hospital of Henan University of Chinese Medicine (Approval No. 2023HL–547) and strictly adhered to the principles of the Declaration of Helsinki. Written informed consent was obtained from all participants or their legal guardians, explicitly authorizing research participation and subsequent data publication (including potentially identifiable information). Fasting peripheral venous blood (5 mL) was collected from CKD patients on the morning following hospital admission and from healthy controls during physical examinations. All whole blood samples were incubated at room temperature for 1 h, centrifuged at 2500 rpm for 15 min, with supernatants transferred to new 2 mL centrifuge tubes and stored at −80°C until analysis.

2.9. Cellular grouping and modeling

Murine renal tubular epithelial cells (TCMK1) were allocated into control and model groups per experimental design. Primary TCMK1 cultures were maintained in DMEM medium supplemented with 10% fetal bovine serum (FBS) under standard conditions (37°C, 5% CO_2). When cells reached 60–70% confluence, the medium was replaced. Control groups maintained standard medium, while model groups were stimulated for 24 hours with complete medium containing 5 ng/mL Transforming growth factor-β 1 (TGF-β1) to induce EMT [34].

2.10. Immunofluorescence staining

Cells from each group were seeded on cover slips and cultured for 24 h, then fixed with 4% paraformaldehyde for 10 min. After permeabilization with 0.5% Triton X-100 for 15 min, samples were blocked with 5% goat serum at room temperature for 1 h. Primary antibodies were incubated overnight at 4°C: α-smooth muscle actin (α-SMA) (1:100), N-cadherin (1:200) and E-cadherin (1:200). Following phosphate-buffered saline (PBS) washes, cells were incubated with FITC-conjugated goat anti-rabbit IgG (1:200) in darkness for 1 h. Nuclei were counterstained with DAPI staining solution for 5 min before fluorescence microscopy imaging.

2.11. RT-qPCR

Cells in logarithmic growth phase were plated at a density of 1×10^6 cells/well in 6-well plates, with triplicate wells per experimental group. Culture medium was aspirated, and 1 ml TRIzol reagent was added per well for complete cell lysis. Frozen serum (−80°C) was thawed at 4°C, and 250 µl aliquots were transferred to 2 ml nuclease-free tubes with 750 µl TRIzol for thorough mixing via pipetting. Total RNA was isolated through phase separation and ethanol precipitation, followed by washing, resuspension, and reverse transcription to synthesize cDNA. cDNA templates were amplified using a two-step PCR protocol: 95°C denaturation (30s) followed by 40 cycles of 95°C (5s) and 60°C (30s). Ct values and melt curves were recorded in triplicate measurements. Relative quantification using $2^{-\Delta\Delta Ct}$ method was performed with primer sequences detailed in Table 2.

2.12. Enzyme linked immunosorbent assay (ELISA)

Human serum CCL2 levels were quantified using an ELISA kit.

2.13. Dual luciferase reporter assays (DLRAs)

Wild-type (CCL2-WT, NEAT1-WT) and mutant (CCL2-MUT, NEAT1-MUT) 3′UTR luciferase reporter plasmids (psi-CHECK2 backbone) were synthesized by General Biol Co., Ltd. China. Lentiviral vectors encoding miR-124-3p-mimic and mimic-negative control (mimic-NC) were constructed by Genepharma Co., Ltd. China. TCMK1 cells were seeded in culture dishes and grown to optimal density. Cells were co-transfected with miR-124-3p-mimic/mimic-NC and dual-luciferase reporter vectors. After 48 hours post-transfection, DLRAs were performed using a luciferase detection kit per manufacturer guidelines.

2.14. Statistical analysis

Normality and homogeneity of variance were assessed for continuous variables. Normally distributed data (Shapiro-Wilk test) were expressed as $\bar{X} \pm S$ and analyzed using t-test (two groups) or one-way ANOVA with SNK-q post hoc testing (multiple groups). Non-normally distributed data were presented as [M (P75-P25)] and compared using Mann-Whitney U test (two groups) or Kruskal-Wallis test with Dunn's correction (multiple groups). Normality and homogeneity of variance were assessed for continuous variables. Categorical variables were expressed as percentages (%) and compared using X^2 test. Bivariate normal distributions were analyzed using Pearson correlation, while Spearman's rank correlation was applied to non-normal/ordinal data. Correlation strength was quantified by coefficient R, where values approaching |1| indicated stronger associations. Receiver Operating Characteristic (ROC) curve analysis evaluated the predictive performance of serum NEAT1, miR-124-3p, and CCL2 levels in clinical risk stratification models. Statistical significance thresholds were set as $^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. All statistical outcomes were verified through triplicate experimental replicates.

3. Results

3.1. Identification of DEGs in GSE137570

Transcriptomic data from the GSE137570 dataset were retrieved from GEO (results in Supplementary Table S1). Differential expression analysis across comparison groups was performed using NetworkAnalyst 3.0 (quality control metrics in Supplementary Figure S1). In Group 1 (Cohort 1, $n = 17$), comparison between CKD progressors ($n = 8$) and non-progressors ($n = 9$) identified 631 DEGs, 514 up, 117 down (Figure 2(A-B)). In Group 2 (Cohort 2, $n = 24$), the eGFR < 30 mL/min/1.73 m^2 subgroup ($n = 9$) versus eGFR ≥ 30 subgroup ($n = 15$) yielded 625 DEGs, 184 up, 441 down (Supplementary Figure S2A, D). Group 3 (Cohort 2, $n = 24$) analysis of eGFR < 60 ($n = 16$) versus ≥60 mL/min/1.73 m^2 ($n = 8$) revealed 331 DEGs, 137 up, 194 down (Supplementary Figure S2B, E). For Group 4 (eGFR < 90 vs. ≥90 mL/min/1.73 m^2), the limited control group size ($n = 3$) compromised

Table 2. Primer sequences.

Gene	Forward primer (5′–3′)	Reverse primer (5′–3′)
mEGF	AGAGCATCTCTCGGATTGACC	CCCGTTAAGGAAAACTCTTAGCA
mVCAN	TGGCCCAGAACGGAAATATCA	ACTAGCCCGGAGTTTGACCAT
mCXCL1	CTGGGATTCACCTCAAGAACATC	CAGGGTCAAGGCAAGCCTC
mMMP7	CTTACCTCGGATCGTAGTGGA	CCCCAACTAACCCTCTTGAAGT
mCCL2	TTAAAAACCTGGATCGGAACCAA	GCATTAGCTTCAGATTTACGGGT
mα-SMA	GTCCCAGACATCAGGGAGTAA	TCGGATACTTCAGCGTCAGGA
mN-cadherin	CGCATTGCCACATACACTCT	TTGGCTGAGGATGGTGTAAG
mE-cadherin	CAGTTCCGAGGTCTACACCTT	TGAATCGGGAGTCTTCCGAAAA
mGAPDH	TGTGTCCGTCGTGGATCTGA	TTGCTGTTGAAGTCGCAGGAG
hmiR-124-3p-RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACTTGGCA	
hmiR-124-3p	CGTAAGGCACGCGGTGAA	AGTGCAGGGTCCGAGGTATT
hNEAT1	TGGCTAGCTCAGGGCTTCAG	TCTCCTTGCCAAGCTTCCTTC
hU6	CTCGCTTCGGCAGCACA	AACGCTTCACGAATTTGCGT

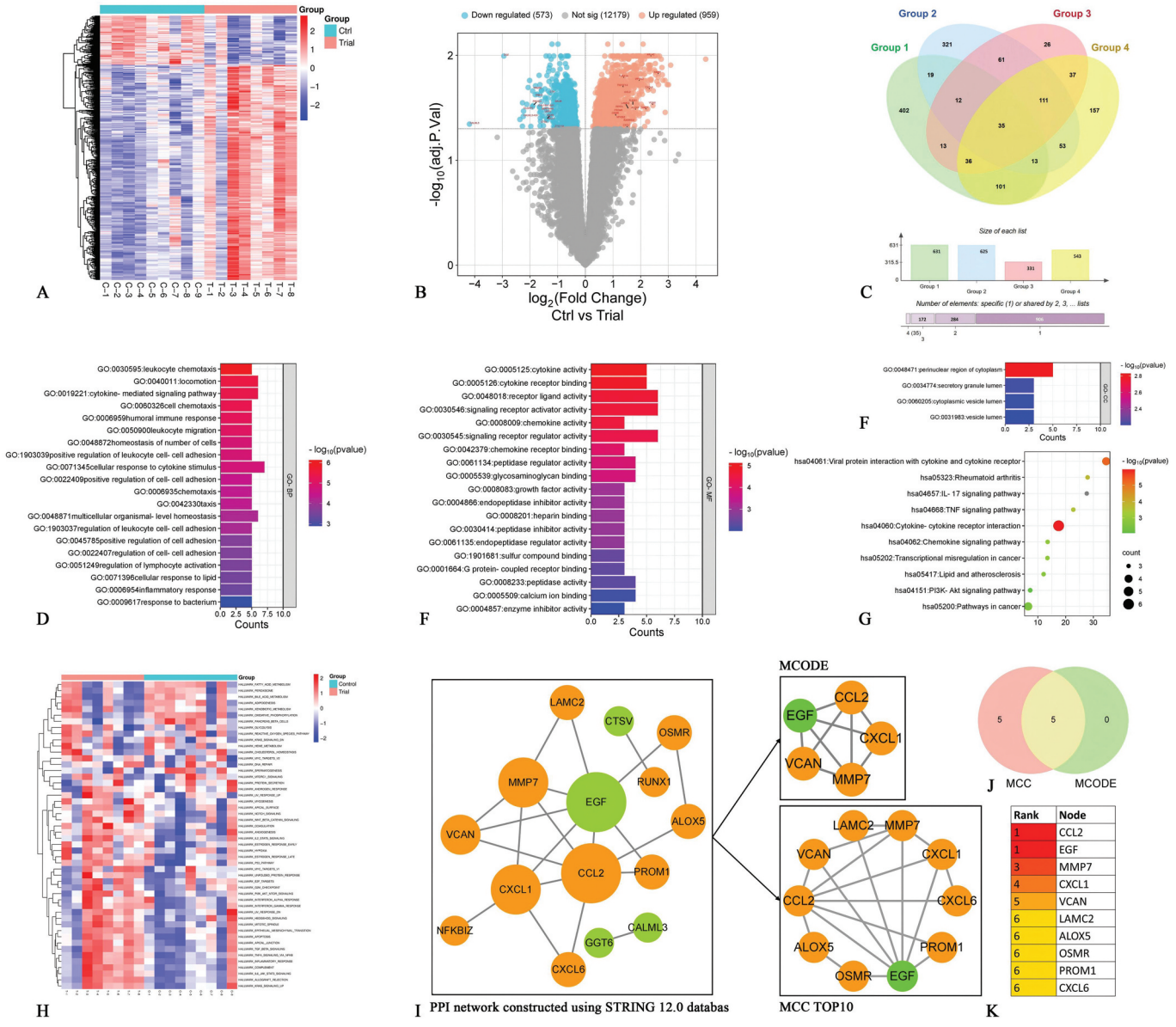

Figure 2. (A) clustering heatmaps of differentially expressed genes (DEGs) across comparison groups: group 1 ($n = 17$): progressors ($n = 8$) vs. Non-progressors ($n = 9$). Rows: DEGs; columns: samples. Color key: red = upregulated genes, blue = downregulated genes. (B) volcano plots of DEGs: red dots: significantly upregulated genes ($\log_2|\text{Fold change (FC)}| \geq 1$, false discovery rate (FDR)-adjusted $p < 0.05$). Blue dots: significantly downregulated genes ($\log_2|\text{FC}| \leq -1$, FDR-adjusted $p < 0.05$). Gray dots: non-significant genes. X-axis: $\log_2$(FC); Y-axis: -$\log_{10}$(adjusted p-value). (C) venn diagram of DEG overlaps across groups: Green: group 1 DEGs (631 genes). Blue: group 2 DEGs (625 genes). Red: group 3 DEGs (331 genes). Yellow: group 4 DEGs (543 genes). Central intersection: 35 common DEGs shared by all groups. (D-G) gene ontology (GO) and Kyoto encyclopedia of genes and genomes (KEGG) enrichment analysis of the 35 common DEGs shared across all four CKD progression groups in GSE137570 dataset. D: GO-Biological process (BP), E: GO-Molecular function (MF), F: GO-Cellular component (CC); G: KEGG pathway analysis. X-axis: gene count per term; Y-axis: -$\log_{10}$(P-value) indicating significance ($p \leq 0.05$). As -$\log_{10}$(P-value) increases, color gradients transition from blue to red in D-F and from green to red in G, corresponding to higher statistical significance. (H) GSVA heatmaps display the differential activity of HALLMARK gene sets between the control and trial groups in groups 1, with each row representing a HALLMARK gene set and each column representing a sample or group. The color gradient indicates the upregulation (red) or downregulation (blue) of gene set activity. (I-K) PPI network construction and Hub gene identification based on the 35 common DEGs shared across all four groups. I: PPI network visualized in Cytoscape: orange nodes represent upregulated DEGs, green nodes represent downregulated DEGs; Hub genes identified by maximal clique centrality (MCC) and molecular complex detection (MCODE) are highlighted. J: venn diagram: overlap of Hub genes identified by MCC (red circle, 10 genes) and MCODE (green circle, 10 genes); the yellow intersection (5 genes: MMP7, VCAN, CXCL1, CCL2, EGF) defines core Hub genes. K: MCC-based node ranking: node color intensity reflects MCC score (darker = higher rank/importance). CCL2 exhibited the highest centrality (darkest node), indicating its pivotal role in CKD progression.

statistical power for robust FDR correction. Given the exploratory nature of this comparison and to mitigate excessive type II errors, differential expression was assessed using nominal P-values ($p < 0.05$) with $|\log_2\text{FC}| > 1$, identifying 631 DEGs, 321 up, 222 down (Supplementary Figure S2C, F). We acknowledge this approach increases false positive risk, and these findings should be considered preliminary (full lists of DEGs are provided in Supplementary Table S2). Collectively, these results demonstrate that our multi-group comparative analysis successfully identified 35 core DEGs (19 up, 16 down) exhibiting high association with CKD progression (Figure 2(C)).

3.2. Enrichment analysis of DEGs

Enrichment analysis of DEGs across comparison groups was performed using NetworkAnalyst, Metascape, and WebGestalt databases.

In Group 1, GO-BP terms were significantly enriched in leukocyte activation and regulation; GO-CC terms localized to the outer plasma membrane and extracellular matrix (ECM); GO-MF involved chemokine/cytokine receptor activity, including G protein-coupled receptor (GPCR) binding. KEGG pathway analysis highlighted cell adhesion molecules and chemokine signaling pathways, indicating coordinated immune cell recruitment and ECM remodeling drive renal interstitial fibrosis (RIF) progression (Supplementary Figure S3). Group 2 exhibited GO-BP enrichment in tumor necrosis factor (TNF) and interleukin-1 (IL-1)-mediated responses; GO-CC enrichment in dendritic structures (e.g., spines); GO-MF dominated by cytokine binding activity. KEGG analysis linked to IL-17 and TNF signaling pathways, reflecting neuroimmune crosstalk that exacerbates tissue injury via inflammatory cascades (Supplementary Figure S4). Group 3 displayed GO-BP enrichment in small molecule and carboxylic acid catabolic processes; GO-CC enrichment in collagen-dense ECM (distinct from Group 1); GO-MF indicating redox enzyme dysregulation. KEGG analysis highlighted immunoglobulin A (IgA)-mediated gut immunity and allograft rejection pathways, suggesting metabolic disturbances potentiate injury through the gut-kidney axis (Supplementary Figure S5). Group 4 featured GO-BP enrichment in B cell-mediated immunity; GO-CC enrichment in stromal compartments (shared features with Group 3) alongside GO-MF enrichment in major histocompatibility complex (MHC) class II binding. KEGG analysis revealed enrichment in autoimmune pathways (e.g., Type 1 diabetes), indicating late-stage adaptive immune dysregulation (Supplementary Figure S6).

Cross-group analysis of shared DEGs identified conserved inflammatory Hub Genes: GO-BP enrichment encompassed cytokine-mediated responses (e.g., IL-17 signaling) and cell motility regulation, suggesting synergy between immune trafficking and inflammatory signaling; GO-CC enrichment in secretory vesicles and luminal compartments (e.g., perinuclear regions) implied exosome-mediated cytokine storage and intercellular transfer; GO-MF centered on cytokine/receptor binding, amplifying inflammatory cascades. Integrated pathway analysis converged on chronic inflammatory circuitry. IL-17/TNF-α signaling fuels innate immunity (Group 2), metabolic dysregulation (Group 3), and autoimmunity (Group 4), establishing trans-group pathological networks. Pan-group enrichment of rheumatoid arthritis pathways suggests targeting shared nodes (e.g., JAK-STAT, IL-17A) may concurrently mitigate fibrosis and autoimmunity, informing cross-stage therapeutic strategies (Figure 2(D-G)).

These results establish a conserved dysregulation of inflammatory and immune responses across progressive stages of CKD, underscoring their central mechanistic role in driving CKD pathogenesis.

3.3. GSVA Analysis

Based on the comparative GSVA analysis between Control and Trial groups across four independent cohorts using MSigDB Hallmark gene sets, we observed the following specific pathway alterations.

In Group 1, the TRIAL group exhibited significant upregulation of fatty acid metabolism and bile acid metabolism, coupled with downregulation of the peroxisome pathway relative to the CONTROL group (Figure 2(H)).

Group 2 demonstrated robust activation of hypoxia signaling and glycolysis in the TRIAL group, while oxidative phosphorylation was suppressed (Supplementary Figure 7A).

Group 3 revealed comprehensive upregulation of inflammation-related pathways in the TRIAL group, including TNFA signaling via NFKB, complement system activation, and IL6-JAK-STAT3 signaling (Supplementary Figure 7B).

Group 4 displayed a distinct dual effect in the TRIAL group: sustained activation of hypoxia signaling alongside potentiated KRAS signaling and EMT, whereas DNA repair pathway activity was suppressed (Supplementary Figure 7C).

These GSVA results collectively delineate distinct molecular signatures across CKD progression stages: an early shift toward metabolic reprogramming (Groups 1–2), a pronounced inflammatory surge in mid-stage disease (Group 3), and a late-stage activation of pro-fibrotic signaling alongside impaired DNA repair (Group 4), highlighting the dynamic evolution of pathogenic pathways during CKD advancement.

3.4. Construction of PPI network

Expression profiles of 35 DEGs across all four experimental groups were analyzed using the STRING 12.0 database to construct a PPI network (Figure 2(I)). This PPI network was imported into Cytoscape for topological assessment and visualization. Cytohubba-based topological analysis identified ten Hub Genes: Laminin Subunit Gamma 2 (LAMC2), Matrix Metalloproteinase 7 (MMP7), Versican (VCAN), C-X-C Motif Chemokine Ligand 1 (CXCL1), C-C Motif Chemokine Ligand 2 (CCL2), CXCL6, Arachidonate 5-Lipoxygenase (ALOX5), Prominin 1 (PROM1), Epidermal Growth Factor (EGF), and Oncostatin M Receptor (OSMR). MCC analysis independently identified MMP7, VCAN, CXCL1, CCL2, and EGF as Hub Genes. The intersection of both methods established MMP7, VCAN, CXCL1, CCL2, and EGF as core Hub Genes, with the CCL2-associated cluster exhibiting the largest network size and highest node ranking (Figure 2(J-K)).

Due to technical heterogeneity in Group 4, a secondary PPI network was constructed using DEGs exclusively from Groups 1–3. Subsequent Cytohubba and MCC analyses identified LAMC2, CCL22, CXCL1, CXCL6, EGF, VCAN, CCL2, and MMP7 as Hub Genes (Supplementary Figure 8). Notably, MMP7, VCAN, CXCL1, CCL2, and EGF were consistently prioritized across both analytical frameworks. The CCL2 cluster again demonstrated maximal size and node centrality, reinforcing its biological significance in disease progression.

Together, these results demonstrate that PPI network analysis and Hub Gene screening established MMP7, VCAN, CXCL1, CCL2, and EGF as key regulators of CKD progression, with CCL2 occupying the topologically central position.

3.5. Screening of Hub-associated miRnas and prediction of target lncRnas

Hub-related miRNAs were screened using MiRWalk and miRTarBase databases. Intersection analysis identified seven EGF-associated miRNAs: hsa-miR-665, hsa-miR-1304-5p, hsa-miR-1827, hsa-miR-4722-5p, hsa-miR-6765-5p, hsa-miR-6840-3p, and hsa-miR-6884-5p. Two VCAN-associated miRNAs were identified: hsa-miR-103a-3p and hsa-miR-107. A single CXCL1-linked miRNA was detected: hsa-miR-196b-5p. Three MMP7-associated miRNAs were identified: hsa-miR-34a-5p, hsa-miR-4697-5p, and hsa-miR-7109-5p. CCL2 was associated with one miRNA: hsa-miR-124-3p (miRNA predictions in Supplementary Table S3).

Further, ENCORI and MiRNET databases were used to predict lncRNAs associated with hub miRNAs. Results showed no relevant lncRNAs were retrieved for hsa-miR-1304-5p, hsa-miR-1827, hsa-miR-4722-5p, hsa-miR-6765-5p, hsa-miR-6840-3p, hsa-miR-4697-5p, or hsa-miR-7109-5p. hsa-miR-665 was ultimately associated with 24 lncRNAs. hsa-miR-6884-5p with 16 lncRNAs. hsa-miR-103a-3p with 16 lncRNAs. hsa-miR-107 with 16 lncRNAs. hsa-miR-196b-5p with 12 lncRNAs. hsa-miR-34a-5p with 13 lncRNAs. hsa-miR-124-3p with 10 lncRNAs. Among these, NEAT1 exhibited the highest CLIP Experiment Number (clipExpNum), Degradation Experiment Number (degraExpNum), and Phylogenetic *P*-value (phylo *P*), indicating the strongest confidence (Figure 3, lncRNA predictions are detailed in Supplementary Table S4).

These results demonstrate the successful prediction of potential regulatory miRNAs for the identified Hub Genes and construction of corresponding ceRNA networks. Notably, the miR-124-3p/CCL2 and NEAT1/miR-124-3p regulatory relationships exhibit high prediction confidence, suggesting that the NEAT1/miR-124-3p/CCL2 axis may constitute a critical ceRNA regulatory pathway.

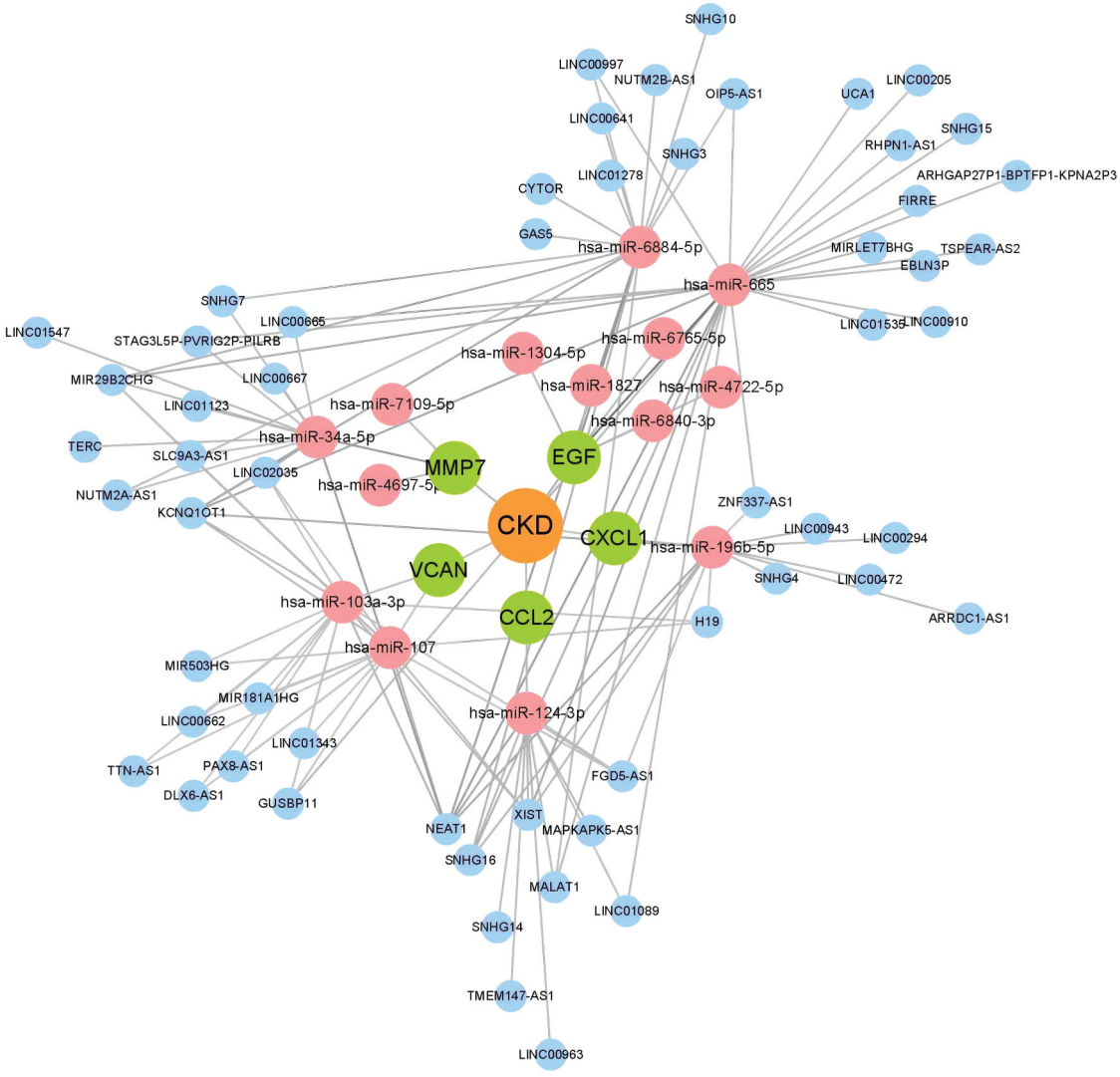

Figure 3. Predicted competing endogenous RNA (ceRNA) network for Hub genes. The network depicts the predicted regulatory interactions. Orange nodes: diseases, green nodes: hub genes (EGF, VCAN, CXCL1, MMP7, CCL2), pink nodes: miRnas predicted to target the hubs, blue nodes: lncRnas predicted to interact with the miRnas. Lines represent predicted interactions (miRNA-mRNA or lncRNA-miRNA).

3.6. Analysis of renal function parameters and Hub gene features in CKD patients from GSE137570 dataset

Analysis of renal function parameters and Hub Genes in the GSE137570 dataset included Cohort 1 with one variable (Renal function progression within 60 months) and Cohort 2 with two variables (eGFR and TIF). CKD patients in trial groups exhibited significantly reduced eGFR levels compared to controls. Elevated TIF scores were observed across all groups except Group 4. Renal tissue analysis demonstrated upregulated VCAN, CXCL1, MMP7, and CCL2 expression, alongside downregulated EGF levels, in CKD groups versus controls (Table 3). Correlation analyses revealed: inverse associations between VCAN/CXCL1/MMP7/CCL2 and eGFR, positive correlations with TIF and disease progression. EGF showed opposing trends (Table 4).

Together, these results demonstrate that expression levels of the identified Hub Genes significantly correlate with renal functional impairment in CKD patients, further supporting their potential as biomarkers for CKD progression.

3.7. Correlation between Hub genes and immune cells

Using the CIBERSORT algorithm in the TIMER 2.0 database, we analyzed differences in immune cell infiltration across renal tissue groups and assessed correlations between differentially infiltrated cells and Hub Genes.

Group 1 exhibited significant differences in B cells naive, T cells CD4$^+$ memory resting, Natural killer cells (NK cells) resting, NK cells activated, and monocytes between CKD and control renal tissues. EGF, VCAN, CXCL1, MMP7, and CCL2 showed moderate correlations with CD4$^+$ memory resting T cells and monocytes, strong correlations with activated NK cells, and weak correlations with resting NK cells (Figure 4(A-B)).

Group 2 displayed significant monocyte differences between CKD and controls, with all Hub Genes showing moderate correlations (Supplementary Figure 9A, D).

Group 3 showed significant monocyte differences; VCAN, CXCL1, MMP7, and CCL2 had moderate correlations, while EGF exhibited weak correlation (Supplementary Figure 9B, E).

Group 4 demonstrated significant differences in CD4$^+$ naive T cells and Regulatory T cells (Tregs), with VCAN, CXCL1, MMP7, and CCL2 showing moderate correlations (Supplementary Figure 9C, F).

These results collectively indicate that the landscape of immune cell infiltration in renal tissue exhibits significant differences across various stages of CKD progression. Moreover, the expression of Hub Genes demonstrates significant correlations with the infiltration abundance of multiple immune cell subtypes, suggesting a cooperative interplay between alterations in the immune microenvironment and Hub Gene expression during CKD progression.

3.8. ROC curve analysis of Hub genes in GSE137570 dataset

ROC curve analysis of Hub Genes in the GSE137570 dataset revealed that in Group 1 (Figure 4(C)), the area under the curve (AUC) values for EGF, VCAN, CXCL1, MMP7, and CCL2

Table 3. eGFR,TIF and hub genes expression (mean ± SD or M[P75-P25]).

Cohort	Grouping	Subgroup	Estimated glomerular filtration rate (eGFR) (min·1.73 m^2)	Tubulointerstitial fibrosis (TIF) (%)	Epidermal Growth Factor (EGF) (Relative expression of transcriptome)	Versican (VCAN)	C-X-C Motif Chemokine Ligand 1 (CXCL1)	Matrix Metalloproteinase 7 (MMP7)	Chemokine (C-C motif) ligand 2 (CCL2)
Cohort 1 ($n = 17$)	Group 1	Progressors ($n = 8$)	–	–	10.66 ± 1.47*	12.67 ± 1.31*	8.90 ± 1.12*	14.78 ± 1.38*	10.91 ± 0.98*
		Non-progressors ($n = 9$)	–	–	13.61 ± 0.94	10.22 ± 1.62	7.39 ± 0.98	12.59 ± 1.33	9.34 ± 0.73
Cohort 2 ($n = 24$)	Group 2	eGFR < 30 ($n = 9$)	24.00 (8.00)*	70.00 (18.75)*	8.03 ± 1.11*	11.58 ± 1.13*	7.54 ± 1.29*	12.92 ± 1.09*	10.00 ± 0.73*
		eGFR ≥30 ($n = 15$)	70.20 ± 34.58	33.67 ± 25.88	10.14 ± 1.29	9.28 ± 1.59	5.42 ± 1.18	10.49 ± 1.53	8.74 ± 0.88
Cohort 2 ($n = 24$)	Group 3	eGFR < 60 ($n = 16$)	29.88 ± 12.36*	57.50 (37.50)*	8.70 ± 1.45*	11.00 ± 1.40*	6.80 ± 1.41*	12.22 ± 1.35*	9.63 ± 0.77*
		eGFR ≥60 ($n = 8$)	95.75 ± 27.21	23.13 ± 22.94	10.64 ± 0.98	8.44 ± 1.26	5.04 ± 1.31	9.75 ± 1.49	8.38 ± 0.98
Cohort 2 ($n = 24$)	Group 4	eGFR < 90 ($n = 21$)	35.00 (34.50)*△	55.00 (48.75)	9.06 ± 1.49*	10.48 ± 1.63*	6.53 ± 1.42*	11.75 ± 1.54*	9.45 ± 0.81*
		eGFR ≥90 ($n = 3$)	126.33 ± 13.05△	5.00 (45.00)	11.31 ± 0.66	7.83 ± 1.40	4.01 ± 0.86	8.96 ± 1.96	7.56 ± 0.85

Table 4. Correlation between Hub gene expression levels and CKD progression, eGFR, and TIF.

Cohort	Index		Epidermal Growth Factor (EGF)	Versican (VCAN)	C-X-C Motif Chemokine Ligand 1 (CXCL1)	Matrix Metalloproteinase 7 (MMP7)	Chemokine (C-C motif) ligand 2 (CCL2)
Cohort 1 (n = 17)	Chronic kidney disease (CKD) progression	R value	−0.7939	0.6736	0.6255	0.6495	0.6736
		P value	0.0006	0.0055	0.0111	0.0079	0.0055
Cohort 2 (n = 24)	Estimated glomerular filtration rate (eGFR)	R value	0.7334	−0.7438	−0.6964	−0.8347	−0.7786
		P value	< 0.0001	< 0.0001	0.0002	< 0.0001	< 0.0001
Cohort 2 (n = 24)	Tubulointerstitial fibrosis (TIF)	R value	−0.4923	0.7544	0.5772	0.6792	0.5915
		P value	0.0145	< 0.0001	0.0031	0.0003	0.0023

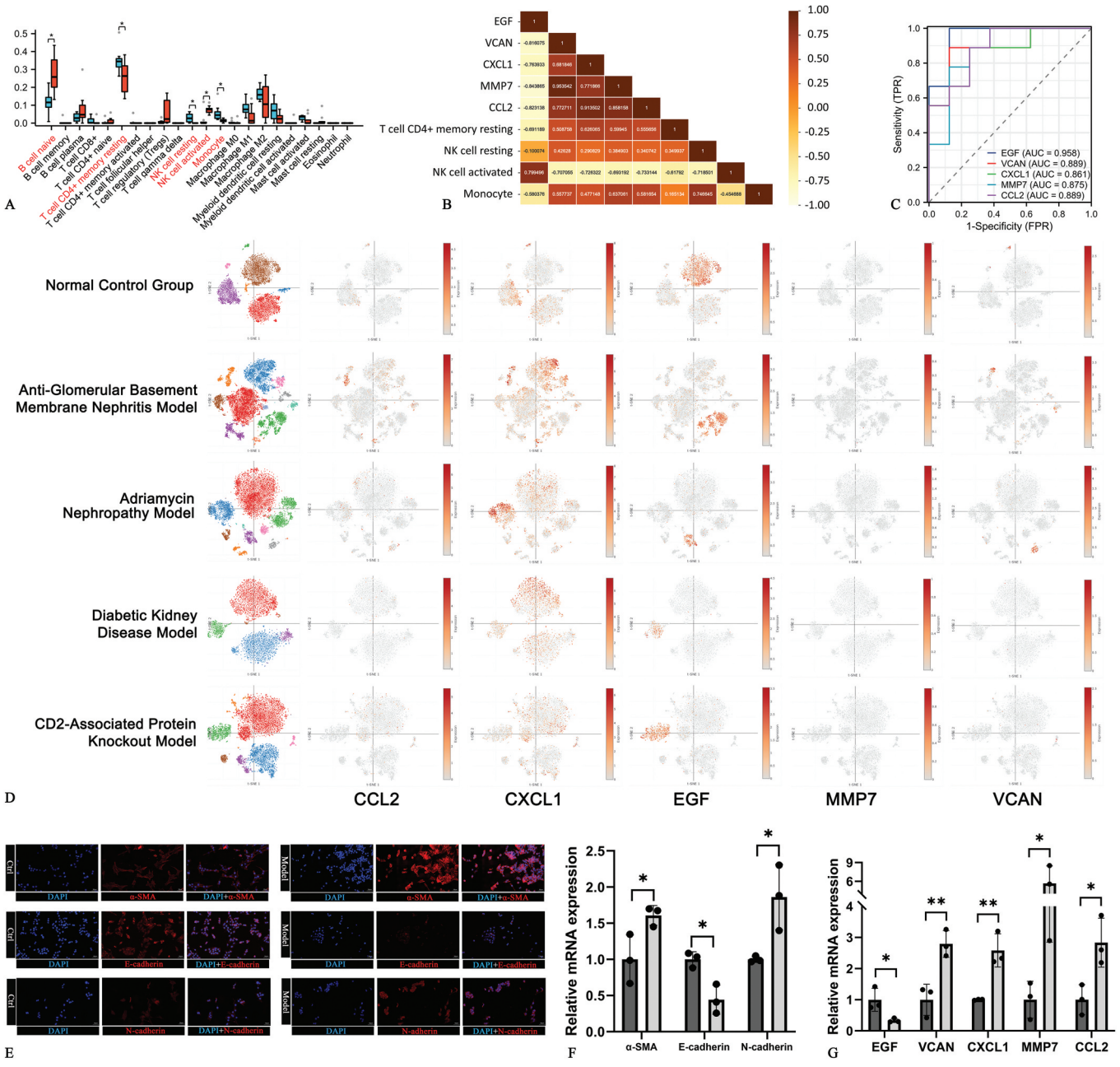

Figure 4. (A) box plots show relative proportions of 22 immune cell subtypes estimated by CIBERSORT in renal tissues (n = 17). Blue boxes: control group (non-progressors). Red boxes: trial group (progressors). Data points represent biological replicates (individual patient samples). Whiskers indicate minimum to maximum values, boxes show interquartile range (IQR), horizontal lines denote medians. Significant immune cell subtypes ($p < 0.05$) are highlighted in red within the box plots and marked with asterisks (*). (B) heatmaps display Spearman correlation coefficients (R values) between significantly different immune cell subtypes and expression levels of the five Hub genes (EGF, VCAN, CXCL1, MMP7, CCL2). Color scale: red indicates positive correlation, yellow indicates negative correlation; darker hues represent stronger correlations (|R| closer to 1). (C) receiver operating characteristic (ROC) curves evaluating the diagnostic performance of EGF, VCAN, CXCL1, MMP7, and CCL2 expression levels in renal tissue. Each colored line represents a different Hub gene. The diagonal black dashed line represents the reference line of no discrimination ((area under the curve (AUC) = 0.5). The AUC value for each Hub gene is shown in the legend. (D) validation of Hub gene expression patterns using public single-cell RNA sequencing data from murine kidney disease models. Visualization of Hub gene (EGF, VCAN, CXCL1, MMP7, CCL2) expression across different cell types within renal tissues of various mouse disease models (normal, anti-GBM nephritis, Adriamycin nephropathy, diabetic kidney disease, CD2AP KO). Color gradient (pale orange to deep red) within each cell type dot represents the expression level of the respective Hub gene; darker red indicates higher expression. The overall cell types present in the tissue are distinguished by color. Red: endothelial cells, blue: arteriolar endothelial cells, Green: immune cells, purple: mesangial cells, orange: myofibroblasts, Brown: glomerular parietal epithelial cells, Pink: podocytes, Gray: smooth muscle cells. (E) Representative immunofluorescence images (630x magnification) showing expression of α-smooth muscle actin (α-SMA), E-Cadherin, and N-Cadherin. Green fluorescence (FITC) indicates target protein localization, blue fluorescence (DAPI) stains nuclei. Model groups show increased α-SMA & N-Cadherin and decreased E-Cadherin signal vs. Control. (F) bar graphs (mean ± standard deviation (SD), n = 3) showing mRNA levels of EMT markers (α-SMA, E-Cadherin, N-Cadherin) by RT-qPCR. Control group: white bars, model group: black bars.* $p < 0.05$, ** $p < 0.01$ vs. control. (G) bar graphs (mean ± SD, n = 3) showing mRNA levels of Hub genes (EGF, VCAN, CXCL1, MMP7, CCL2) by RT-qPCR. Control group: black bars, model group: white black bars. * $p < 0.05$, ** $p < 0.01$ vs. control.

were 0.958 (95% Confidence interval (CI): 0.867–1.000), 0.889 (95% CI:0.723–1.000), 0.861 (95% CI: 0.682–1.000), 0.875 (95% CI: 0.694–1.000), and 0.889 (95% CI: 0.733–1.000), respectively, in Group 2 (Supplementary Figure 9 G), the corresponding AUC values were 0.881 (95% CI: 0.747–1.000), 0.889 (95% CI: 0.756–1.000), 0.904 (95% CI: 0.763–1.000), 0.919 (95% CI: 0.808–1.000), and 0.859 (95% CI: 0.707–1.000); Group 3 demonstrated AUC values of 0.883 (95% CI: 0.730–1.000), 0.922 (95% CI: 0.817–1.000), 0.820 (95% CI: 0.636–1.000), 0.906 (95% CI: 0.789–1.000), and 0.836 (95% CI: 0.657–1.000) for these biomarkers (Supplementary Figure 9 H); In Group 4 (Supplementary Figure 9I), the AUC values reached 0.952 (95% CI: 0.865–1.000), 0.921 (95% CI: 0.753–1.000), 0.952 (95% CI: 0.846–1.000), 0.873 (95% CI: 0.614–1.000), and 0.968 (95% CI: 0.893–1.000) for these respective markers. These findings suggest that EGF, VCAN, CXCL1, MMP7, and CCL2 demonstrate robust predictive capabilities for the progression of renal injury in CKD.

These results collectively confirm that EGF, VCAN, CXCL1, MMP7, and CCL2 all demonstrate superior diagnostic performance across distinct groups within the GSE137570 dataset. This indicates their potential utility as biomarkers for predicting the progression of CKD.

3.9. Validation of Hub gene expression

Validation of the selected Hub Genes was performed using single-cell RNA sequencing data from the Single Cell Portal database and further confirmed through in vitro experiments (Figure 4(D-G)).

Immunofluorescence detection revealed that α-SMA, a transmembrane protein, exhibited positive signals primarily localized to the cell membrane and cytoplasmic regions. Positive expression of E-cadherin and N-cadherin was mainly observed at the cell membrane and within cell junctions and adherens junctions. Compared to the control group, the TGF-β1-induced model group showed significantly increased fluorescence intensity for α-SMA and N-cadherin, while the E-cadherin fluorescence signal was markedly attenuated (Supplementary Figure 10).

RT-qPCR results further validated these model characteristics: the model group exhibited significantly upregulated mRNA expression of α-SMA and N-cadherin and significantly downregulated mRNA expression of E-cadherin, confirming the successful establishment of the renal tubular epithelial cell transdifferentiation model (The RT-qPCR data are presented in Supplementary Table S5).

EGF validation results demonstrated significantly lower renal tissue EGF expression in the Anti-Glomerular Basement Membrane Nephritis, Adriamycin Nephropathy, Diabetic Kidney Disease, and CD2AP Knockout models compared to the normal control group. In vitro experiments similarly showed significantly reduced EGF mRNA expression.

VCAN validation results indicated significantly elevated VCAN expression in the Anti-GBM and Adriamycin Nephropathy models, but no significant changes in the Diabetic Kidney Disease and CD2AP models. In vitro experiments revealed significantly upregulated VCAN mRNA expression in the model group.

CXCL1 validation results showed significantly increased expression across all kidney disease models. In vitro experiments demonstrated significantly upregulated CXCL1 mRNA expression in the model group.

MMP7 validation results showed consistently low expression levels without significant changes in all in vivo models. However, in vitro experiments demonstrated significant upregulation of MMP7 mRNA in the model group. To further investigate MMP7 expression in the RIF model, we employed a unilateral ureteral obstruction (UUO) mouse model to mimic the RIF microenvironment. Western blot analysis revealed markedly increased MMP7 protein expression in renal tissues of the model group, consistent with the in vitro results (Supplementary Figure S11).

CCL2 validation results showed significantly increased expression only in the Anti-GBM model, with no significant changes in other models. In vitro experiments indicated significantly upregulated CCL2 mRNA expression in the model group.

By combining single-cell sequencing with in vitro experimental validation, the expression patterns of Hub Genes were consistent with predictions from previous analyses.Given that CCL2 exhibited the highest node centrality within the PPI network, and its predicted sole associated miR-124-3p and upstream lncRNA NEAT1 demonstrated the highest confidence within the ceRNA network, subsequent investigations in this study will focus on CCL2 and its regulatory pathway NEAT1/miR-124-3p for in-depth mechanistic validation and exploration of clinical significance.

3.10. Targeting interaction of miR-124-3p with NEAT1 and CCL2

To validate the specific regulatory axis between NEAT1/miR-124-3p and miR-124-3p/CCL2, dual-luciferase reporter vectors containing wild-type (WT) or mutant (MUT) miR-124-3p binding sites in NEAT1 and CCL2 were constructed. DLRAs revealed that miR-124-3p mimic transfection significantly suppressed the relative luciferase activity of NEAT1-WT and CCL2-WT compared to other co-transfection groups, confirming direct targeting interactions between miR-124-3p and NEAT1/CCL2 (Figure 5(A-C)).

These results collectively confirm that DLRA validation demonstrated that miR-124-3p directly binds to the 3′UTR regions of both NEAT1 and CCL2. This provides direct experimental evidence supporting the existence of a ceRNA regulatory axis wherein NEAT1 modulates CCL2 expression by sequestering miR-124-3p.

3.11. Basic information of included samples

This study addresses the clinical need for noninvasive biomarkers with high sensitivity and specificity in early-stage CKD diagnosis by investigating the diagnostic potential of serum non-coding RNAs (ncRNAs). Building on CKD renal tissue transcriptomic features identified in the GSE137570 dataset, we further validated the expression patterns of NEAT1, miR-124-3p, and CCL2 mRNA in serum. The study cohort comprised 64 patients with CKD G2-G4 and 20 healthy controls, with serum marker expression measured via RT-qPCR and correlated with renal function indices such as eGFR and BUN.

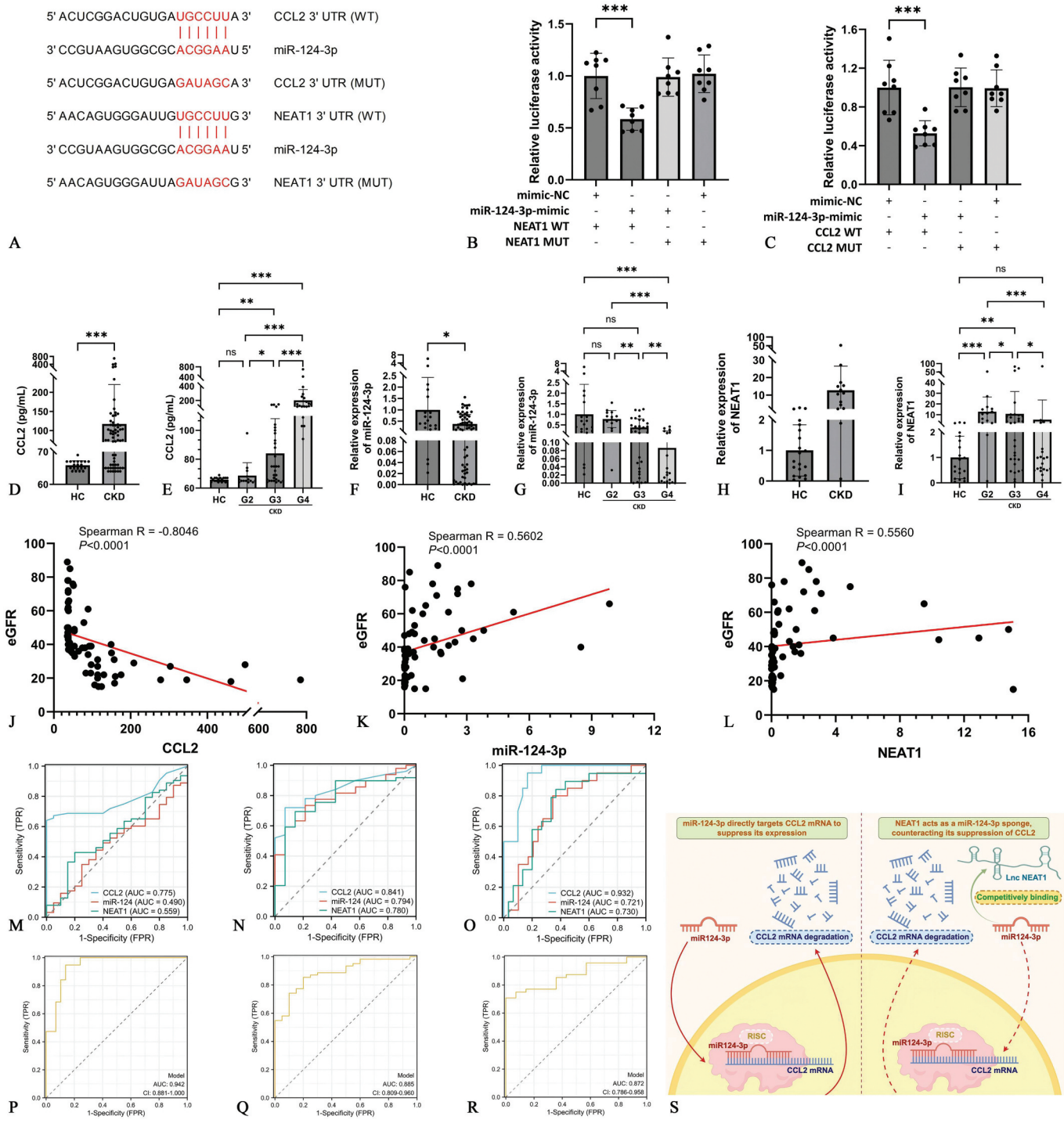

Figure 5. (A) schematic representation of the pmirGLO dual-luciferase vector constructs. WT 3′UTR sequences of NEAT1 or CCL2 containing the predicted miR-124-3p binding site, and corresponding MUT sequences with binding site disruption, were cloned downstream of the renilla luciferase gene. (B-C) relative luciferase activity in TCMK1 cells co-transfected with miR-124-3p mimic or mimic negative control (NC) and the indicated reporter plasmids (NEAT1-WT, NEAT1-MUT, CCL2-WT, CCL2-MUT). Data are presented as mean ± SD ($n = 3$). *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. (D,E) bar graphs display serum levels of CCL2. (F,G) bar graphs display serum levels of miR-124-3p. (H,I) bar graphs display serum levels of NEAT1, as mean ± SD with individual data points representing biological replicates (healthy controls $n = 20$; CKD G2 $n = 14$; CKD G3 $n = 30$; CKD G4 $n = 20$). *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. (J-L) scatter plots depict the correlation between expression levels of CCL2, miR-124-3p, and NEAT1 (y-axis) and eGFR (x-axis), with each data point representing an independent biological sample. (M-R) ROC curves where colored lines represent biomarkers and black dashed line denotes no discrimination (AUC = 0.5); AUC values with 95% CI indicate diagnostic accuracy (0.5–0.7 = low, 0.7–0.9 = moderate, > 0.9 = high); panels M, N, O show individual biomarkers distinguishing: NORMAL vs. CKD ($n = 84$: 20 controls + 64 CKD), CKD G2 vs. G3&G4 ($n = 64$: 14 G2 + 50 G3/G4), and CKD G3 vs. G4 ($n = 50$: 30 G3 + 20 G4); panels P, Q, R show combined biomarkers for the same group comparisons respectively. (S) schematic depiction summarizing the proposed mechanism. miR-124-3p directly targets CCL2 mRNA to suppress its expression. NEAT1, as a miR-124-3p sponge, competitively binds to miR-124-3p, downregulating its inhibitory effect on CCL2 transcription. This results in the upregulation of CCL2 expression, which subsequently exacerbates kidney inflammation infiltration and fibrosis processes. The dashed lines indicate suppression or reduction in the expression or activity of the targeted molecules(this figure was drawn by Figdraw (https://www.figdraw.com/) with the ID: PTITY0abdd.).

Baseline analysis demonstrated no significant differences in demographic characteristics (gender, age, and body mass index (BMI)) between CKD patients and healthy controls, ensuring intergroup comparability. Comparative analysis of serum biochemical parameters showed significantly elevated BUN and Scr levels coupled with markedly reduced eGFR in CKD patients versus healthy controls. Progressive deterioration of renal function parameters was observed across CKD stages: Stage G3 patients exhibited significantly lower eGFR and higher cystatin C (Cys C) levels compared to stage G2, while stage G4 demonstrated further elevations in Scr, BUN, Cys C, and 24-hour urinary total protein (24hUTP) relative to stage G3. Results are presented in Table 5.

These results collectively demonstrate the successful enrollment of eligible CKD patients and healthy controls with comparable baseline characteristics. Furthermore, renal function indices in CKD patients exhibited a graded deterioration corresponding to disease stage progression. This establishes a reliable clinical sample foundation for subsequent serum biomarker analysis.

3.12. Comparative analysis of serum CCL2, miR-124-3p, and NEAT1 levels

Serum levels of CCL2 (quantified via ELISA) and miR-124-3p/NEAT1 (measured by RT-qPCR) were analyzed in 20 healthy controls and 64 CKD patients. Compared to controls, CKD patients exhibited significantly elevated CCL2 levels, whereas miR-124-3p and NEAT1 showed no significant alterations. Intriguingly, progressive increases in Scr correlated with rising CCL2 levels and concurrent declines in miR-124-3p and NEAT1 expression across CKD stages (Figure 5(D-I)). Correlation analyses revealed: (1) inverse associations between CCL2 and miR-124-3p/NEAT1; (2) positive miR-124-3p/NEAT1 correlation; (3) strong links between CCL2/miR-124-3p and renal markers (Figure 5J-L), Supplementary Figure 12). ROC curve analysis demonstrated CCL2's robust diagnostic efficacy for both CKD onset and progression. While miR-124-3p/NEAT1 showed limited predictive value for disease initiation, they exhibited improved prognostic utility in tracking CKD advancement (Figure 5(M-R)).

These results collectively demonstrate that serum levels of CCL2, miR-124-3p, and NEAT1 exhibit dynamic CKD patients. Furthermore, their expression levels are interrelated and show significant correlations with renal function impairment indices. More significantly, they demonstrate favorable diagnostic value in distinguishing different stages of CKD progression. This suggests their potential utility as noninvasive biomarkers for monitoring CKD progression in clinical practice.

4. Discussion

This study identified five Hub Genes EGF, VCAN, CXCL1, MMP7, and CCL2 associated with CKD progression through bioinformatics analysis and PPI network screening. These genes are closely linked to renal function impairment and may hold significant clinical diagnostic value for renal injury, suggesting their potential for clinical translation. It is noteworthy, however, that Group 4 exhibited a significant sample size imbalance. Consequently, DEG detection relied solely on p-values. While this approach allowed the preliminary identification of 543 candidate genes, the absence of FDR correction increases the risk of false positives. These results should therefore be interpreted as hypothesis-generating rather than definitive. The decision to use p-values reflects a deliberate trade-off between statistical rigor and biological exploration. We emphasize that this exception does not apply to other analyses within this study, where FDR correction was consistently applied. Crucially, a key validation step strengthens our confidence in the Hub Genes: MMP7, VCAN, CXCL1, CCL2, and EGF consistently appeared in the hub screening results for Groups 1–3. This convergence indicates that these genes exhibit strong CKD associations, and the statistical methodological compromise in Group 4 did not compromise the identification of these core targets.

Furthermore, through Hub Gene screening, ceRNA network prediction, and DLRA validation, we identified the NEAT1/miR-124-3p/CCL2 axis as a potential ceRNA regulatory pathway in CKD. NcRNAs represent a class of transcripts that lack protein-coding potential. Ubiquitously expressed across cell types, ncRNAs are genomically transcribed yet function directly at the RNA level without translation [35,36]. LncRNAs, defined as ncRNAs exceeding 200 nucleotides, regulate gene expression through epigenetic, transcriptional, and post-transcriptional mechanisms. They modulate target genes either by direct binding or via recruitment of transcription factors to activate or repress expression [37]. Conversely, miRNAs constitute short (~22 nt) ncRNAs detectable in multiple biofluids (serum, plasma, urine). Their expression profiles reflect systemic responses to pathophysiological stressors, mediating cellular processes including activation, proliferation, transformation, apoptosis, lipid metabolism, and renal fibrotic progression [38]. As early-response molecules to renal injury, miRNAs demonstrate particular utility as biomarkers and represent

Table 5. Comparison of clinical and biochemical parameters (mean ± SD or M[P75-P25]).

Grouping		HC ($n = 20$)	CKD G2 ($n = 14$)	CKD G3 ($n = 30$)	CKD G4 ($n = 20$)
Gender	[male/rate (%)]	9.00 (45.00)	11.00 (78.60)	15.00 (50.00)	13.00 (65.00)
Age	(year)	49.85 ± 13.53	50.86 ± 11.71	62.5 (21.50)	56.45 ± 12.05
Body mass index (BMI)	(kg/m^2)	25.13 ± 2.969	24.6 (6.97)	22.73 ± 2.882	25.92 ± 3.759
Serum creatinine (Scr)	(μmol/L)	67.75 (7.37)	110.03 ± 18.03[a]	118.05 (57.68)[a]	343.80 ± 115.20[abc]
Estimated glomerular filtration rate (eGFR)	[ml/(min·1.73 m^2)]	106.00 ± 15.18	71.36 ± 9.27[a]	40.47 ± 6.17[ab]	21.50 ± 4.44[abc]
	(mmol/L)	5.03 ± 2.361	6.86 (2.47)	10.63 (4.50)[a]	16.68 (7.13)[abc]
Uric acid (UA)	(μmol/L)	–	351.09 ± 103.94	374.72 ± 86.36	442.23 ± 139.53[b]
Cystatin C (Cys C)	(mg/L)	–	1.10 (0.32)	1.58 ± 0.37[b]	2.43 ± 0.43[bc]
24-hour urinary total protein (24 h UTP)	(g/24 h)	–	2.26 (2.06)	1.17 (2.08)	4.80 (6.43)[bc]
Total protein (TP)	(g/L)	–	65.88 ± 10.06	63.10 (21.25)	60.70 ± 9.90

a complementary diagnostic modality for CKD detection [39–41].

Mounting evidence has elucidated the multidimensional regulatory mechanisms of NEAT1 in kidney diseases. NEAT1 mediates pivotal roles in renal injury pathogenesis by dynamically regulating inflammation, fibrosis, and cell death pathways. During AKI, NEAT1 activates NLR family pyrin domain containing 3 (NLRP3) inflammasome assembly through the receptor for activated C kinase 1 (Rack1)-mediated Toll-like receptor 4/nuclear factor kappa B (TLR4/NF-κB) signaling axis, thereby amplifying inflammatory responses. Concurrently, it promotes renal tubular epithelial cell apoptosis via p53-dependent transcriptional regulation [42,43]. Functioning as a ceRNA, NEAT1 sequesters miR-22-3p, consequently attenuating its inhibitory effect on the CXCL12/NF-κB pathway. NEAT1 further exacerbates oxidative stress and apoptotic cascades through the miR-330-5p/Forkhead box O3 (FOXO3) axis, establishing a synergistic inflammation-tissue injury amplification network [44,45]. During chronic disease progression, NEAT1 directly binds miR-129-5p, relieving its suppression of Fas-associated protein with death domain/caspase-8/caspase-3 (FADD/CASP-8/3) apoptotic signaling. This cascade activates programmed cell death in renal tubular cells and propels interstitial fibrosis [43]. In established fibrosis, NEAT1 represses miR-31, attenuating its negative regulation of the Ras homolog family member A/Rho-associated coiled-coil containing protein kinase (RhoA/ROCK) pathway, leading to significant upregulation of fibrosis markers including connective tissue growth factor (CTGF). Within metabolically stressed microenvironments, NEAT1 dynamically modulates EMT through the miR-204/SRY-box transcription factor 4 (SOX4) axis, thereby promoting fibroblast activation and ECM deposition [21].

miR-124-3p is a microRNA that exerts significant regulatory functions in kidney pathophysiology. It mediates critical biological processes across diverse renal pathologies including diabetic nephropathy, nephrolithiasis, AKI, lupus nephritis, and renal interstitial fibrosis [46]. By modulating downstream target gene expression, miR-124-3p regulates key pathological mechanisms such as inflammation [47], oxidative stress [48], autophagy dysregulation [49], and ferroptosis [50], thereby influencing renal cellular homeostasis and disease pathogenesis. Research demonstrates that miR-124-3p directly targets signal transducer and activator of transcription 3 (STAT3), inhibiting its pro-inflammatory signaling and promoting macrophage polarization toward the anti-inflammatory M2 phenotype, thereby attenuating inflammation-driven fibrosis [51]. In podocytes, miR-124-3p exerts anti-inflammatory and anti-fibrotic effects while suppressing apoptosis through inhibition of the calpain-1 (CAPN1)/β-catenin signaling pathway in diabetic nephropathy models [52]. Furthermore, miR-124-3p targets and inhibits Toll-like receptor 4 (TLR4), blocking activation of the TLR4/NF-κB signaling pathway. This suppression reduces production of downstream fibrosis-associated factors including tumor TNF-α and IL-6, consequently attenuating renal fibrogenesis [53]. In the UUO model, miR-124-3p upregulation occurs through competitive binding with lncRNAs (e.g., metastasis-associated lung adenocarcinoma transcript 1 (MALAT1), HOX transcript antisense RNA (HOTAIR)); it subsequently directly inhibits integrin subunit beta 1 (ITGB1) or indirectly inhibits jagged canonical Notch ligand 1 (JAG1), thereby blocking the ITGB1 and JAG1/Notch1 signaling axes. This dual inhibition effectively suppresses EMT and excessive ECM deposition, mitigating fibrosis progression [54,55]. Collectively, these studies establish miR-124-3p as a multi-target regulator modulating multiple pathways, with anti-fibrotic actions encompassing inflammation suppression, EMT regulation, ECM deposition reduction, and normalization of aberrant cellular proliferation and apoptosis.

Upon renal tissue injury, the innate immune system initiates a cascade of inflammatory mediators including cytokines, chemokines, and reactive oxygen species (ROS). These mediators accumulate within the renal interstitium, triggering fibroblast proliferation, activation, and transdifferentiation into myofibroblasts. This process promotes excessive synthesis of collagen fibers and ECM components, ultimately replacing functional renal parenchyma with fibrotic scar tissue. Chemokines – secreted signaling proteins characterized by their ability to induce directional chemotaxis – are classified into four subfamilies (CXC, CC, CX3C, and XC) based on conserved cysteine residue motifs. Chemokines mediate biological functions by binding G protein-coupled transmembrane receptors (GPCRs), playing crucial roles in leukocyte migration, cellular proliferation, pathogen defense, and the pathogenesis of inflammation and cancer within tissue microenvironments [56]. CCL2, a CC-subfamily member, exhibits potent chemotactic activity for monocytes and basophils. It recruits CCR2-expressing leukocytes to sites of tissue injury or infection, amplifying inflammatory cascades [57]. During renal damage, tubule epithelial cells release CCL2, attracting monocytes, T lymphocytes, dendritic cells, and other immune cells to the injury site. These recruited cells promote tissue remodeling while participating in both initial injury and maladaptive repair processes [58,59]. Research indicates that CCL2 undergoes significant dysregulation during early renal injury. Its detectable levels in serum and urine exhibit favorable stability, suggesting utility for noninvasive monitoring of disease activity and therapeutic response [60].

Based on the preceding evidence, we hypothesize that renal tissue injury induces elevated NEAT1 expression, which competitively binds miR-124-3p, thereby attenuating miR-124-3p-mediated suppression of CCL2. This cascade amplifies inflammatory responses and exacerbates renal functional impairment (Figure 5(S)). However, the specific involvement of the NEAT1/miR-124-3p/CCL2 axis in CKD pathogenesis remains unreported. To validate CCL2 and its regulators NEAT1, miR-124-3p, we conducted a pilot clinical study revealing dysregulated serum expression of all three molecules in CKD patients, with levels correlating with renal functional decline – suggesting potential utility for early diagnosis and progression monitoring. Intriguingly, while all biomarkers demonstrated diagnostic promise, their expression patterns diverged from expected trends. Although accumulated evidence consistently reports significant NEAT1 upregulation in CKD renal tissues (including diabetic kidney disease and RIF models), where elevated levels correlate positively with disease progression and NEAT1 inhibition attenuates fibrosis in

experimental models, our clinical findings revealed no uniform serum NEAT1/miR-124-3p expression differences between controls and the overall CKD cohort. Further analysis demonstrated significantly elevated serum NEAT1 and miR-124-3p in stage G2 CKD patients versus healthy controls, while CCL2 remained comparable to controls. As CKD progressed from G2 to G4, serum NEAT1 and miR-124-3p levels progressively decreased, whereas CCL2 expression significantly increased. The miR-124-3p decline likely reduces its repression of CCL2, aligning with canonical miRNA regulatory mechanisms. Paradoxically, the observed positive correlation between NEAT1 and miR-124-3p expression contradicts the anticipated inverse relationship predicted by the ceRNA hypothesis.

To reconcile this paradox, we propose the following non-mutually exclusive mechanisms:

(1) Phase-Specific Compensatory Surge: The concurrent elevation of serum NEAT1 and miR-124-3p in early-stage CKD (G2) may represent a transient adaptive response. miR-124-3p upregulation could constitute a physiological effort to mitigate initial renal injury through anti-inflammatory/fibrotic actions, while NEAT1 induction reflects a counter-regulatory stress signal. This compensatory interplay may transiently override canonical ceRNA dynamics during early pathogenesis.
(2) Multi-Layered Transcriptional Dominance: NEAT1 expression is governed by regulatory hierarchies extending beyond miRNA sponging, including: Transcription factor activation (e.g., hypoxia-inducible factors under renal hypoxia) [61]; Epigenetic reprogramming (e.g., METTL14-mediated m6A methylation modulating NEAT1 stability) [62]; Chromatin remodeling at super-enhancer loci during CKD progression [63]; Post-transcriptional stabilization via RNA-binding proteins [64]. These context-dependent pathways may supersede ceRNA regulation in early CKD, decoupling NEAT1 from miR-124-3p.
(3) Technical and Biological Confounders: Sample heterogeneity: Variability in CKD etiology (e.g., diabetic vs. hypertensive nephropathy) may drive distinct biomarker signatures. Preanalytical factors: Differences in RNA stability in circulation, circadian rhythms, or sample processing protocols could quantitatively bias measurements. Compartmentalization: Serum levels may not faithfully mirror intrarenal molecular crosstalk due to tissue-blood barrier effects or cell-type-specific expression.

However, it is important to acknowledge the limitation imposed by the restricted size of the clinical cohort in this pilot investigation. The modest sample size, particularly within early-stage CKD G2 subgroups, may compromise statistical power for detecting subtle biomarker dynamics and increase vulnerability to type II statistical errors. Furthermore, the cross-sectional design precludes evaluation of longitudinal biomarker trajectories. These limitations may thus impair the generalizability of our findings concerning the paradoxical elevation of NEAT1/miR-124-3p in CKD stage G2. To address this limitation, future large-scale, multi-center prospective studies incorporating serial sampling are warranted to validate these preliminary observations and delineate the trajectory of the NEAT1/miR-124-3p/CCL2 axis across CKD stages.

While prior studies have independently implicated NEAT1, miR-124-3p, and CCL2 in RIF, the functional significance of the NEAT1/miR-124-3p/CCL2 axis in RIF remains incompletely defined. Through integrated bioinformatic prediction, DLRA, and clinical sample validation, our study establishes the critical involvement of this axis in CKD progression. These findings highlight its dual utility as both a diagnostic biomarker and a promising therapeutic target. Targeted modulation of this axis – for instance, via inhibition of NEAT1 expression or function – may yield novel therapeutic strategies to attenuate CKD progression. Antisense oligonucleotide (ASO) technology, a validated approach for lncRNA targeting [65], could be leveraged here. ASOs designed against NEAT1 transcripts can effectively suppress NEAT1 stability and function through RNase H-mediated degradation or steric blockade [66]. Building on our results, developing NEAT1-directed ASO therapies may alleviate its molecular sponge function toward miR-124-3p. This restoration of miR-124-3p activity would enhance its repression of the downstream pro-inflammatory and pro-fibrotic factor CCL2, thereby mitigating renal inflammatory infiltration and fibrosis. Although not experimentally validated in this study, such a strategy constitutes a feasible translational research avenue emerging from our discovery. To further elucidate the axis's role in CKD, future work will employ in vivo and in vitro RIF models. This will include generating tissue-specific knockout models (e.g., tubular epithelial cell-selective Neat1 knockout mice) and establishing NEAT1-transcriptionally silenced cell lines. Subsequent functional rescue experiments will delineate the regulatory mechanisms of the Neat1/miR-124-3p/CCL2 axis in RIF pathogenesis.

5. Conclusion

In summary, this study identified and validated EGF, VCAN, CXCL1, and CCL2 as novel biomarkers with diagnostic and therapeutic potential for CKD, while additionally highlighting the NEAT1/miR-124-3p/CCL2 axis as a promising diagnostic and therapeutic target. However, this study has limitations: First, bioinformatic analysis relied on the GSE137570 dataset with its modest sample size, potentially introducing bias. Second, in vitro experimental validation cannot fully replicate the complexity of the in vivo microenvironment or CKD's multifaceted progression. Third, the limited clinical cohort may restrict generalizability and introduce sampling bias. Finally, the precise molecular mechanisms of the NEAT1/miR-124-3p/CCL2 axis in CKD pathophysiology remain incompletely defined. Future research should prioritize large-scale, multi-center prospective clinical trials for validation, complemented by mechanistic studies using in vivo and refined in vitro models to fully dissect this axis's role in CKD pathogenesis. Such efforts will advance novel diagnostic and therapeutic strategies for CKD management.

Acknowledgments

The authors would like to thank the Student Research Committee, Henan University of Chinese Medicine, for their helpful advice.

The authors acknowledge the use of DeepSeek-R1 (v1.0), an AI-assisted language model developed by DeepSeek, for proofreading and grammatical refinement of the manuscript text. The authors supervised and reviewed all AI-generated content to ensure scientific accuracy and adherence to the original data.

Disclosure statement

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

No writing assistance was utilized in the production of this manuscript.

Peer reviewers on this manuscript have no relevant financial or other relationships to disclose.

Funding

This manuscript was funded by National Natural Science Foundation of China (82374381). Youth Program of the Natural Science Foundation of Henan Province (242300421522). Special Project of Henan National Clinical Research Base of Chinese Medicine (2022JDZX109). Special Project on Chinese Medicine Research in Henan Province (2023ZY1002). Henan Provincial Postdoctoral Research Project (HN2025078). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author contributions

Guanting Chen: Conceptualization, Methodology, Formal analysis, Investigation, Writing – Original Draft, Writing – Review & Editing

Linqi Zhang: Supervision, Project administration, Funding acquisition, Writing – Review & Editing (Corresponding author)

Yaoxian Wang: Resources, Data curation, Writing – Review & Editing (Corresponding author)

Jianfeng Wang: Methodology, Validation, Writing – Review & Editing

Kang Yang: Investigation, Data curation, Writing – Review & Editing

Xixi Wang: Formal analysis, Visualization, Writing – Original Draft

Xu Chen: Visualization, Formal analysis, Writing – Original Draft

All authors have read and approved the final manuscript.

Ethical Declaration

This study was approved by the Ethics Committee of the First Affiliated Hospital of Henan University of Chinese Medicine (No. 2021HL–109–01) and complied with the Declaration of Helsinki. Written informed consent was obtained from all participants or their legal representatives.

Data availability statement

The publicly available datasets GSE137570 and SCP975 are authorized for research analysis under the open-access policies of their respective repositories. In compliance with the NCBI GEO Data Usage Policy and the Single Cell Portal Open Access Agreement, downloading and analyzing these datasets do not require additional permission from the data owners.

The data that support the findings of this study are available from the corresponding author, Linqi Zhang, upon reasonable request.

References

Papers of special note have been highlighted as either of interest (•) or of considerable interest (••) to readers.

1. Ying M, Shao X, Qin H, et al. Disease burden and epidemiological trends of chronic kidney disease at the global, regional, National levels from 1990 to 2019. Nephron. 2024;148(2):113–123. doi: 10.1159/000534071
2. Foreman KJ, Marquez N, Dolgert A, et al. Forecasting life expectancy, years of life lost, and all-cause and cause-specific mortality for 250 causes of death: reference and alternative scenarios for 2016–40 for 195 countries and territories. Lancet. 2018;392(10159):2052–2090. doi: 10.1016/S0140-6736(18)31694-5
3. Di X, Li Y, Wei J, et al. Targeting fibrosis: from molecular mechanisms to advanced therapies. Adv Sci (Weinh). 2025;12(3):e2410416. doi: 10.1002/advs.202410416
4. Miguel V, Tituaña J, Herrero JI, et al. Renal tubule Cpt1a overexpression protects from kidney fibrosis by restoring mitochondrial homeostasis. J Clin Invest. 2021;131(5):e140695. doi: 10.1172/JCI140695
5. Fogo AB, Harris RC. Crosstalk between glomeruli and tubules. Nat Rev Nephrol. 2025;21(3):189–199. doi: 10.1038/s41581-024-00907-0
6. Rocha KB, Soares VA, Viero RM. The role of myofibroblasts and interstitial fibrosis in the progression of membranous nephropathy. Ren Fail. 2004;26(4):445–451. doi: 10.1081/jdi-200026764
7. Alexopoulos E, Papagianni A, Papadimitriou M. Is membranous nephropathy only a glomerular disease? Ren Fail. 1998;20(1):1–6. doi: 10.3109/08860229809045084
8. Huang R, Fu P, Ma L. Kidney fibrosis: from mechanisms to therapeutic medicines. Signal Transduct Target Ther. 2023;8(1):129. doi: 10.1038/s41392-023-01379-7
9. Niculae A, Gherghina ME, Peride I, et al. Pathway from acute kidney injury to chronic kidney disease: molecules involved in renal fibrosis. Int J Mol Sci. 2023;24(18):14019. doi: 10.3390/ijms241814019
10. Stevens PE, Ahmed SB, Carrero JJKidney Disease: Improving Global Outcomes (KDIGO) CKD Work Group. KDIGO 2024 clinical practice guideline for the evaluation and management of chronic kidney disease. Kidney Int. 2024;105(4):S117–S314. doi: 10.1016/j.kint.2023.10.018
11. Nastase MV, Zeng-Brouwers J, Wygrecka M, et al. Targeting renal fibrosis: mechanisms and drug delivery systems. Adv Drug Deliv Rev. 2018;129:295–307. doi: 10.1016/j.addr.2017.12.019
12. Rende U, Guller A, Goldys EM, et al. Diagnostic and prognostic biomarkers for tubulointerstitial fibrosis. J Physiol. 2023;601(14):2801–2826. doi: 10.1113/JP284289
13. Ma B, Wang S, Wu W, et al. Mechanisms of circRna/lncrna-miRnainteractions and applications in disease and drug research. Biomed Pharmacother. 2023;162:114672. doi: 10.1016/j.biopha.2023.114672
14. Ghafouri-Fard S, Abak A, Talebi SF, et al. Role of miRNA and lncRnas in organ fibrosis and aging. Biomed Pharmacother. 2021;143:112132. doi: 10.1016/j.biopha.2021.112132
15. Yu S, Li Y, Lu X, et al. The regulatory role of miRNA and lncRNA on autophagy in diabetic nephropathy. Cell Signal. 2024;118:111144. doi: 10.1016/j.cellsig.2024.111144
16. Salmena L, Poliseno L, Tay Y, et al. A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? Cell. 2011;146(3):353–358. doi: 10.1016/j.cell.2011.07.014
17. Li K, Wang Z. lncRNA NEAT1: key player in neurodegenerative diseases. Ageing Res Rev. 2023;86:101878. doi: 10.1016/j.arr.2023.101878
18. Jiang X. The mechanisms and therapeutic potential of long noncoding RNA NEAT1 in fibrosis. Clin Exp Med. 2023;23(7):3339–3347. doi: 10.1007/s10238-023-01191-1
•• **Note: Instruction: Systematically elaborate on the mechanism of NEAT1 in fibrosis to provide core theoretical support for the ceRNA axis in this study.**
19. Huang S, Xu Y, Ge X, et al. Long noncoding RNA NEAT1 accelerates the proliferation and fibrosis in diabetic nephropathy through activating Akt/mTOR signaling pathway. J Cell Physiol. 2019;234(7):11200–11207. doi: 10.1002/jcp.27770

20. Yang YL, Xue M, Jia YJ, et al. Long noncoding RNA NEAT1 is involved in the protective effect of Klotho on renal tubular epithelial cells in diabetic kidney disease through the ERK1/2 signaling pathway. Exp Mol Med. 2020;52(2):266–280. doi: 10.1038/s12276-020-0381-5
21. Ahmad Wadaan M, Ahmad M, Hussein Hamzah I. Long non-coding RNA NEAT1 inhibits high glucose-induced EMT and renal fibrogenesis in human embryonic kidney 293 cells via regulating miR-204/SOX4 axis. Cell Mol Biol (noisy-le-Grand). 2023;69(10):189–194. doi: 10.14715/cmb/2023.69.10.27
22. Yan Q, Hu Q, Li G, et al. NEAT1 regulates calcium oxalate Crystal-induced renal tubular oxidative injury via miR-130/IRF1. Antioxid Redox Signal. 2023;38(10–12):731–746. doi: 10.1089/ars.2022.0008
23. Chen Y, Huang C, Duan ZB, et al. LncRNA NEAT1 accelerates renal fibrosis progression via targeting miR-31 and modulating RhoA/ROCK signal pathway. Am J Physiol Cell Physiol. 2023;324(2):C292–C306. doi: 10.1152/ajpcell.00382.2021
•• **Note: Directly validate that NEAT1 regulates renal fibrosis via miRNAs to support the molecular mechanism proposed in this study.**
24. Qin Z, Wang PY, Su DF, et al. miRNA-124 in immune system and immune disorders. Front Immunol. 2016;7:406. doi: 10.3389/fimmu.2016.00406
25. Ghafouri-Fard S, Shoorei H, Bahroudi Z, et al. An update on the role of miR-124 in the pathogenesis of human disorders. Biomed Pharmacother. 2021;135:111198. doi: 10.1016/j.biopha.2020.111198
26. Yuan Q, Tang B, Zhang C. Signaling pathways of chronic kidney diseases, implications for therapeutics. Signal Transduct Target Ther. 2022;7(1):182. doi: 10.1038/s41392-022-01036-5
27. Wang Q, Liu F, Li Y, et al. Choroid plexus CCL2–CCR2 signaling orchestrates macrophage recruitment and cerebrospinal fluid hypersecretion in hydrocephalus. Acta Pharm Sin B. 2024;14(10):4544–4559. doi: 10.1016/j.apsb.2024.06.020
28. Chang FC, Liu CH, Luo AJ, et al. Angiopoietin-2 inhibition attenuates kidney fibrosis by hindering chemokine C-C motif ligand 2 expression and apoptosis of endothelial cells. Kidney Int. 2022;102(4):780–797. doi: 10.1016/j.kint.2022.06.026
•• **Note: Confirm the central role of CCL2 in renal fibrosis to support its potential as a therapeutic target.**
29. Clough E, Barrett T. The gene expression omnibus database. Methods Mol Biol. 2016;1418:93–110. doi: 10.1007/978-1-4939-3578-9_5
30. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics. 2013;14:7. doi: 10.1186/1471-2105-14-7
31. Udayasuryan B, Zhou Z, Ahmad RN, et al. Fusobacterium nucleatum infection modulates the transcriptome and epigenome of HCT116 colorectal cancer cells in an oxygen-dependent manner. Commun Biol. 2024;7(1):551. doi: 10.1038/s42003-024-06201-w
32. Chung JJ, Goldstein L, Chen YJ, et al. Single-cell transcriptome profiling of the kidney glomerulus identifies key cell types and reactions to injury. J Am Soc Nephrol. 2020;31(10):2341–2354. doi: 10.1681/ASN.2020020220
33. Rossing P, Caramori ML, Chan JCNKidney Disease: Improving Global Outcomes (KDIGO) Diabetes Work Group. KDIGO 2022 clinical practice guideline for diabetes management in chronic kidney disease. Kidney Int. 2022;102(5):S1–S127. doi: 10.1016/j.kint.2022.06.008
34. Peng Z, Guo HY, Li YQ, et al. The Smad3-dependent microRNA let-7i-5p promoted renal fibrosis in mice with unilateral ureteral obstruction. Front Physiol. 2022;13:937878. doi: 10.3389/fphys.2022.937878
35. Chen LL, Kim VN. Small and long non-coding RNAs: past, present, and future. Cell. 2024 Nov 14;187(23):6451–6485. doi: 10.1016/j.cell.2024.10.024
36. Nemeth K, Bayraktar R, Ferracin M, et al. Non-coding RNAs in disease: from mechanisms to therapeutics. Nat Rev Genet. 2024;25(3):211–232. doi: 10.1038/s41576-023-00662-1
37. Mercer TR, Munro T, Mattick JS. The potential of long noncoding RNA therapies. Trends Pharmacol Sci. 2022;43(4):269–280. doi: 10.1016/j.tips.2022.01.008
38. Lv W, Fan F, Wang Y, et al. Therapeutic potential of microRNAs for the treatment of renal fibrosis and CKD. Physiol Genomics. 2018;50(1):20–34. doi: 10.1152/physiolgenomics.00039.2017
• **Note: Demonstrate the therapeutic potential of miRNAs in CKD to provide a theoretical framework for this study.**
39. Mahtal N, Lenoir O, Tinel C, et al. MicroRNAs in kidney injury and disease. Nat Rev Nephrol. 2022;18(10):643–662. doi: 10.1038/s41581-022-00608-6
40. Aomatsu A, Kaneko S, Yanai K, et al. MicroRNA expression profiling in acute kidney injury. Transl Res. 2022;244:1–31. doi: 10.1016/j.trsl.2021.11.010
41. Cerqueira DM, Tayeb M, Ho J. MicroRNAs in kidney development and disease. JCI Insight. 2022;7(9):e158277. doi: 10.1172/jci.insight.158277
42. Xue R, Yiu WH, Chan KW, et al. Long non-coding RNA NEAT1, NOD-Like receptor family protein 3 inflammasome, and acute kidney injury. J Am Soc Nephrol. 2024;35(8):998–1015. doi: 10.1681/ASN.0000000000000362
43. Ma T, Li H, Liu H, et al. Neat1 promotes acute kidney injury to chronic kidney disease by facilitating tubular epithelial cells apoptosis via sequestering miR-129-5p. Mol Ther. 2022;30(10):3313–3332. doi: 10.1016/j.ymthe.2022.05.019
44. Zhou Y, Wang Y, Li Q, et al. Downregulation of lncRNA NEAT1 alleviates sepsis-induced acute kidney injury. Cent Eur J Immunol. 2022;47(1):8–19. doi: 10.5114/ceji.2022.115628
45. Gong Y, Dong X, Xu J, et al. LncRNA NEAT1 knockdown ameliorates LPS-induced human kidney injury by mediating the miR-330-5p/FOXO3 axis. Int Urol Nephrol. 2022;54(10):2683–2694. doi: 10.1007/s11255-022-03179-4
46. Chen G, Wang Y, Zhang L, et al. Research progress on miR-124-3p in the field of kidney disease. BMC Nephrol. 2024;25(1):252. doi: 10.1186/s12882-024-03688-7
•• **Note: Comprehensively review the role of miR-124-3p in kidney disease to provide critical background for the regulatory axis investigated in this study.**
47. Zhang S, Dong D, Zhang Y, et al. miR-124-3p relieves allergic rhinitis by inhibiting dipeptidyl peptidase-4. Int Immunopharmacol. 2021;101(Pt B):108279. doi: 10.1016/j.intimp.2021.108279
48. Guo X, Li C, Wang Y, et al. Long non-coding RNA nuclear paraspeckle assembly transcript 1 downregulation protects lens epithelial cells from oxidative stress-induced apoptosis by regulating the microRNA-124-3p/death-associated protein kinase 1 axis in age-related cataract. Int Ophthalmol. 2023;43(9):3413–3424. doi: 10.1007/s10792-023-02749-4
49. Wu L, Tian X, Zuo H, et al. miR-124-3p delivered by exosomes from heme oxygenase-1 modified bone marrow mesenchymal stem cells inhibits ferroptosis to attenuate ischemia-reperfusion injury in steatotic grafts. J Nanobiotechnology. 2022;20(1):196. doi: 10.1186/s12951-022-01407-8
50. Zhao J, Wang Y, Wang D, et al. MiR-124-3p attenuates brain microvascular endothelial cell injury in vitro by promoting autophagy. Histol Histopathol. 2022;37(2):159–168. doi: 10.14670/HH-18-406
51. Sun Y, Li Q, Gui H, et al. MicroRNA-124 mediates the cholinergic anti-inflammatory action through inhibiting the production of pro-inflammatory cytokines. Cell Res. 2013;23(11):1270–1283. doi: 10.1038/cr.2013.116
52. Zhao N, Du L, Ma Y, et al. LncRNA NEAT1/microRNA-124 regulates cell viability, inflammation and fibrosis in high-glucose-treated mesangial cells. Exp Ther Med. 2022;24(2):507. doi: 10.3892/etm.2022.11434
53. Zhang S, Xu L, Liang R, et al. Baicalin suppresses renal fibrosis through microRNA-124/TLR4/NF-κB axis in streptozotocin-induced diabetic nephropathy mice and high glucose-treated human proximal tubule epithelial cells. J Physiol Biochem. 2020;76(3):407–416. doi: 10.1007/s13105-020-00747-z
•• **Note: Verify that miR-124-3p exerts anti-fibrotic effects via inflammatory pathways to support the regulatory mechanism proposed in this study.**
54. Chen F, Hu Y, Xie Y, et al. Total glucosides of paeony alleviate cell apoptosis and inflammation by targeting the long noncoding RNA

XIST/MicroRNA-124-3p/ITGB1 axis in renal Ischemia/Reperfusion injury. Mediators Inflamm. 2020 [cited 2020 Nov 24];2020:1–13. doi: 10.1155/2020/8869511
55. Xia W, Chen X, Zhu Z, et al. Knockdown of lncRNA MALAT1 attenuates renal interstitial fibrosis through miR-124-3p/ITGB1 axis. Sci rep-Uk. 2023;13(1):18076. doi: 10.1038/s41598-023-45188-y
56. Ozga AJ, Chow MT, Luster AD. Chemokines and the immune response to cancer. Immunity. 2021;54(5):859–874. doi: 10.1016/j.immuni.2021.01.012
57. Lin Z, Shi JL, Chen M, et al. CCL2: an important cytokine in normal and pathological pregnancies: a review. Front Immunol. 2023;13:1053457. doi: 10.3389/fimmu.2022.1053457
58. Lv W, Booz GW, Wang Y, et al. Inflammation and renal fibrosis: recent developments on key signaling molecules as potential therapeutic targets. Eur J Pharmacol. 2018;820:65–76. doi: 10.1016/j.ejphar.2017.12.016
59. Haller H, Bertram A, Nadrowitz F, et al. Monocyte chemoattractant protein-1 and the kidney. Curr Opin Nephrol Hypertens. 2016;25(1):42–49. doi: 10.1097/MNH.0000000000000186
60. Moloi MW, Rusch JA, Omar F, et al. Urinary MCP-1 and TWEAK as non-invasive markers of disease activity and treatment response in patients with lupus nephritis in South Africa. Int Urol Nephrol. 2021;53(9):1865–1873. doi: 10.1007/s11255-020-02780-9
61. Ji A, Li H, Fu X, et al. Long non-coding RNA NEAT1 induced by BHLHE40 activates Wnt/β-catenin signaling and potentiates colorectal cancer progression. Cell Div. 2024;19(1):25. doi: 10.1186/s13008-024-00129-7
62. Liu TC, Li HX, Wan YX, et al. METTL14-mediated upregulation of lncRNA HOTAIR represses PP1α expression by promoting H3K4me1 demethylation in oxycodone-treated mice. CNS Neurosci Ther. 2024;30(7):e14830. doi: 10.1111/cns.14830
63. Ma X, Mei S, Wuyun Q, et al. Super-enhancer-driven LncRNA PPARα-seRNA exacerbates glucolipid metabolism and diabetic cardiomyopathy via recruiting KDM4B. Mol Metab. 2024;86:101978. doi: 10.1016/j.molmet.2024.101978
64. Hämmerle M, Gutschner T, Uckelmann H, et al. Posttranscriptional destabilization of the liver-specific long noncoding RNA HULC by the IGF2 mRNA-binding protein 1 (IGF2BP1). Hepatology. 2013;58(5):1703–1712. doi: 10.1002/hep.26537
65. Momtazmanesh S, Rezaei N. Long non-coding RNAs in diagnosis, treatment, prognosis, and progression of glioma: a state-of-the-art review. Front Oncol. 2021;11:712786. doi: 10.3389/fonc.2021.712786
- **Note: Provide lncRNA-targeting therapeutic strategies (e.g. ASO technology) to provide a basis for the therapeutic recommendations arising from this study.**
66. Naveed A, Cooper JA, Li R, et al. NEAT1 polyA-modulating antisense oligonucleotides reveal opposing functions for both long non-coding RNA isoforms in neuroblastoma. Cell Mol Life Sci. 2021;78(5):2213–2230. doi: 10.1007/s00018-020-03632-6