Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Identification of biomarkers and potential therapeutic targets for cadmium-associated nephrolithiasis using bioinformatics

Zhen Zhang
First Affiliated Hospital of Anhui Medical University
Xiaoma Zhang
First Affiliated Hospital of Anhui Medical University
Lei Zhou
First Affiliated Hospital of Anhui Medical University
Zongsan Chen
First Affiliated Hospital of Anhui Medical University
Zongyao Hao
haozongyao@ahmu.edu.cn
First Affiliated Hospital of Anhui Medical University

Research Article

Keywords: cadmium exposure, kidney stones, bioinformatics analysis, IL-11, fibroblasts

Posted Date: November 23rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11311463/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Purpose

Cadmium exposure is closely associated with kidney stone development, yet its mechanisms in nephrolithiasis remain unclear. This study aimed to identify high-confidence candidate molecular targets in cadmium-induced kidney stone pathogenesis via integrated multi-omics bioinformatics analysis.

Methods

Cadmium-related targets, kidney stone targets, DEGs from GSE73680 and key WGCNA modules were intersected to obtain candidate targets. PPI network, GO enrichment and multiple machine learning algorithms were used to prioritize 8 candidates. For IL11, we performed differential expression, GSEA, immune and single-cell localization analyses, together with virtual gene perturbation in fibroblasts and tetracycline–IL11 molecular docking.

Results

Eight candidate targets were identified: IL11, TNFRSF11B, SPP1, AKR1B1, H19, PDZK1, SERPINA1 and CCK. Six proteins formed a PPI network enriched in cytokine activity. All machine learning methods ranked IL11 as the core target.IL11 expression was significantly upregulated in the plaque mucosal group; genes associated with samples exhibiting high IL-11 expression were primarily enriched in pathways such as lysosomes, focal adhesion, and cell adhesion molecules, and were also associated with the proportion of certain immune cells. Single-cell analysis localized IL11 mainly to fibroblasts, which may regulate inflammation and matrix remodeling.

Conclusion

Eight high-confidence targets for cadmium-related kidney stones were screened. IL11 is recognized as the core target, offering new insights into the molecular mechanisms of cadmium-induced nephrolithiasis.

1. Introduction

Kidney stones are one of the most common conditions in urology. Globally, their incidence and recurrence rates continue to rise; in China, the prevalence of kidney stones reaches 7.54%, with a recurrence rate of approximately 50% over 5–10 years. This condition seriously jeopardizes patients' health and imposes a heavy medical and socioeconomic burden[1, 2]. Calcium oxalate stones account for 65.9%–80% of all kidney stones and are the most common type of stone in clinical practice[2, 3]. The Randall's plaque theory is the mainstream theory regarding the pathogenesis of calcium oxalate stones:

calcium phosphate crystals deposit within the renal papillary stroma to form Randall's plaques; these plaques break through the renal papillary epithelium outward, and upon contact with urine, further induce the nucleation, aggregation, and retention of calcium oxalate crystals, ultimately developing into mature kidney stones[3]. Previous studies have largely focused on intrinsic genetic factors as well as extrinsic factors such as diet, lifestyle, and infection. Despite numerous studies, the pathogenesis of kidney stones remains unclear. With the advancement of environmental medicine, the role of environmental heavy metal exposure in the development and progression of kidney stones has gradually garnered attention.

Cadmium is a highly toxic heavy metal pollutant widely present in the environment. It can enter the human body through soil, water, food, tobacco, and other sources, and is preferentially accumulated in the kidneys, causing irreversible damage to the renal tubules[4–6]. Multiple epidemiological observational studies have confirmed that elevated blood and urinary cadmium levels are significantly associated with an increased risk of kidney stones, and this effect is more pronounced in populations living in cadmium-contaminated areas[7]; Cadmium exposure can impair renal tubular reabsorption function and increase the risk of stone formation[7, 8]. However, current epidemiological evidence only demonstrates a correlation between the two; the specific molecular mechanisms by which cadmium exposure promotes kidney stones remain unclear.

Therefore, this study screened a set of candidate molecules by conducting a joint analysis of cadmium toxicity-related targets, kidney stone disease targets, differentially expressed genes from the GSE73680 dataset, and key gene modules identified by WGCNA. Focusing on the four categories of high-confidence candidate targets identified through this screening, we conducted protein–protein interaction (PPI) network analysis and GO functional enrichment analysis, and used machine learning to prioritize the candidate molecules. Building on this, with IL-11 as the central focus, we conducted molecular expression profile analysis, immune infiltration correlation analysis, and KEGG-GSEA enrichment analysis; utilized the GSE231569 dataset to achieve single-cell localization; performed computational gene perturbation simulations in fibroblasts; and completed Tetracycline-IL-11 molecular docking. Through this multi-layered, exploratory bioinformatics study, this research aims to identify potential candidate molecules and pathways that can be validated in subsequent independent cohorts and verified through in vivo and in vitro functional experiments, thereby providing a theoretical framework for elucidating the potential molecular mechanisms by which environmental cadmium exposure contributes to the pathogenesis of kidney stones.

2. Materials and Methods

2.1 Screening for Cadmium-Related Targets

The Comparative Toxicogenomics Database (CTD) is a toxicogenomics database that integrates relationships among environmental chemicals, genes/proteins, diseases, and phenotypes, and can be used to analyze potential molecular links between environmental exposure and disease [9]. Using

“cadmium” as the search term, we retrieved gene/protein interaction information related to cadmium exposure, and then standardized the gene names and removed duplicates from the results.

2.2 Screening for Kidney Stone-Related Targets

The Human Gene Compendium (GeneCards) database is an integrated database of human gene information that summarizes all known information about each gene, including its function, expression, sequence, and associated diseases. Online Mendelian Inheritance in Man (OMIM) is a database of human genetic disorders that focuses on documenting the associations between genes and genetic diseases and provides detailed clinical descriptions and pathogenesis. A search was conducted using the keywords “nephrolithiasis,” “kidney stone,” and “ renal calculi”[10, 11], and the results were standardized by gene name and deduplicated.

2.3 Transcriptome Dataset and Differential Expression Analysis

The disease transcriptome data were obtained from the GSE73680 dataset (platform GPL17077) in the GEO database. This dataset originated from a study by Taguchi et al. on the gene expression profiles of Randall’s plaque and normal renal papillary tissue in patients with calcium oxalate stone formation [12]. In this study, samples related to Randall’s plaque were defined as the “Plaque mucosa” group, while samples of non-plaque normal renal papilla tissue were defined as the “Normal mucosa” group, to conduct transcriptomic differential analysis between locally affected tissue and non-plaque renal papilla tissue. This dataset comprises a total of 62 samples, including 33 from the “Normal mucosa” group and 29 from the “Plaque mucosa” group. The GEO series matrix indicates that this dataset provides normalized signals.

2.4 WGCNA and Intersection Analysis

WGCNA was performed using the GSE73680 expression matrix to identify co-expression modules associated with the renal calculus-related tissue phenotype. The WGCNA method followed the framework for weighted gene co-expression network analysis[13]. In this study, the soft threshold power was set to 6.

2.5 PPI Networks and Functional Enrichment Analysis

Protein-protein interaction (PPI) relationships were obtained from the STRING database; H19 (lncRNA) was excluded and not included in the PPI network construction. The interaction network results were imported into Cytoscape software for visualization. Enrichment analysis was performed based on the GO and KEGG databases; the GO and KEGG results were interpreted according to the Gene Ontology and KEGG databases, respectively[14, 15]. The $clusterProfi \leq r$ R package was used to perform the enrichment analysis, and the results—including GeneRatio, P-value, p.adjust, q-value, and involved genes—were reported[16, 17].

2.6 Prioritization of Candidate Genes Based on Machine Learning

The machine learning models were constructed using the support vector machine (SVM), LASSO regression, and random forest algorithm frameworks[18–20]. Model performance was evaluated using the area under the ROC curve (AUC) and cross-validation; interpretation of ROC-AUC results followed standard practices for classical ROC analysis[21, 22]. Given the limited sample size of GSE73680, this study did not include an independent external validation cohort.

2.7 Analysis of the Correlation Between IL-11 Expression and Immune Infiltration

Expression comparisons were performed for IL-11. Based on the expression matrix, the differences in IL-11 expression between the “Plaque mucosa” group (n = 29) and the “Normal mucosa” group (n = 33) were recalculated. Statistical tests were performed using the Welch t-test and the Wilcoxon–Mann–Whitney test, respectively, to verify the robustness of the expression difference results. Using the LM22 feature matrix, the abundance proportions of 22 types of immune cells in the samples were estimated via the CIBERSORT algorithm to further analyze the correlation between IL-11 expression levels and the proportions of various immune cell types[23].

2.8 IL-11-Related GSEA

Using sample IL-11 expression levels as phenotypic labels, we performed a KEGG-GSEA gene set enrichment analysis to identify the transcriptional pathway characteristics of samples with high IL-11 expression. The GSEA analysis workflow and interpretation of gene set enrichment were based on classic GSEA and MSigDB literature[24, 25].

2.9 Single-Cell Localization Analysis

The single-cell analysis was based on the GSE231569 human renal papilla mononuclear/single-cell transcriptomic dataset. This dataset was derived from the human renal papilla mononuclear transcriptomic and spatial transcriptomic atlas constructed by Canela et al., which was used to analyze the cellular composition, immune damage, and matrix remodeling characteristics of the renal papilla in patients with calcium oxalate stones and healthy subjects[26]. Following quality control, normalization, dimensionality reduction, clustering, and cell type annotation, a total of 44,962 cells/cell nuclei were included, comprising 27,829 from the Control group and 17,133 from the Stone group. Single-cell data analysis followed the Seurat methodological framework[27–29], and data processing and visualization were performed within the R ecosystem. Cell annotation was based on classic renal cell marker genes and cluster markers, with a focus on analyzing the expression distribution of IL-11 across different cell types and describing trends in the cell-type-specific log2FC of IL-11 between the Stone and Control groups for each cell type.

2.10 Virtual Perturbation Analysis Based on scTenifoldKnk

Given that IL11 is primarily localized in fibroblasts, this study conducted a computational perturbation analysis using scTenifoldKnk in fibroblasts. scTenifoldKnk predicts gene function virtually based on perturbations of single-cell gene regulatory networks[30].

2.11 Virtual Screening and Molecular Docking

Virtual screening and molecular docking were performed using IL-11 as the target, and a total of 1,522 ligands were evaluated. The virtual screening was conducted using the AutoDock Vina command line. AutoDock Vina and its updated versions are widely used as molecular docking tools to predict protein-ligand binding conformations and binding energies[31, 32]. The ligand ranked first overall was ZINC000150338755; In this study, based on the primary mechanism of action of IL-11 and interpretability, tetracycline was selected as the mechanism-priority candidate ligand for illustration and discussion.

2.12 Animals and experimental design

Male C57BL/6J mice aged 6–8 weeks, purchased from the Animal Center of Anhui Medical University, were housed under specific pathogen-free conditions with free access to standard chow and water. After acclimatization, the mice were randomly assigned to three groups: the saline control group (Saline), glyoxylate group (Gly), and glyoxylate plus cadmium group (Gly + Cd). To establish cadmium exposure, mice in the Gly + Cd group were provided drinking water containing cadmium chloride ($CdCl_2$, 25 mg/L) for 8 weeks, whereas mice in the Saline and Gly groups received normal drinking water. During the final 7 days of cadmium exposure, mice in the Gly and Gly + Cd groups received intraperitoneal injections of glyoxylate (100 mg/kg) once daily to induce renal calcium oxalate (CaOx) crystal deposition, while mice in the Saline group received an equivalent volume of normal saline. Twenty-four hours after the final glyoxylate injection, mice were anesthetized via intraperitoneal injection of 1% pentobarbital sodium at a dose of 50 mg/kg. After adequate anesthesia was confirmed, animals were euthanized, and kidney tissues were immediately collected. Portions of the kidneys were fixed in 4% paraformaldehyde, embedded in paraffin, and sectioned for H&E staining, PAS staining, assessment of renal CaOx crystal deposition, and immunofluorescence analysis. All animal procedures were conducted in accordance with institutional guidelines and were approved by the Institutional Animal Care and Use Committee of Anhui Medical University.

2.13 H&E staining

Paraffin-embedded kidney tissues were sectioned and subjected to routine hematoxylin and eosin (H&E) staining. Briefly, sections were dewaxed in xylene and hydrated using a graded series of ethanol solutions. Following rinsing, the sections were stained with hematoxylin, differentiated, and blued to visualize the nuclei. Eosin staining was subsequently performed to visualize the cytoplasmic and extracellular components. After staining, the sections were dehydrated with graded ethanol, cleared with xylene, mounted using neutral resin, and observed under a light microscope.

2.14 Detection of renal CaOx crystals

Renal calcium oxalate (CaOx) crystal deposition was evaluated in H&E-stained kidney sections. The sections were first examined by conventional light microscopy to identify refractile crystal deposits within the renal tissue. The same sections were then observed under polarized light microscopy (Zeiss, Oberkochen, Germany) to visualize the characteristic birefringence of the deposited crystals. Representative images were acquired for subsequent assessment of renal crystal deposition.

2.15 Tubular injury determination

Renal tubular damage was assessed by periodic acid–Schiff (PAS) staining using a commercially available kit (Beyotime, Shanghai, China). Paraffin-embedded kidney sections were dewaxed, rehydrated, and exposed to 0.5% periodic acid for 10 min at room temperature. After washing with distilled water, the sections were treated with Schiff's reagent for 20 min in the dark and subsequently washed with sulfite solution. Nuclear counterstaining was performed with hematoxylin for 2 min. The stained sections were then processed according to the manufacturer's recommended protocol and examined for histopathological alterations in renal tubules.

2.16 Immunofluorescence staining

Immunofluorescence staining was performed to evaluate the localization of IL11 and PDGFRα in renal tissues. Paraffin-embedded kidney sections were deparaffinized, rehydrated, and subjected to antigen retrieval. After washing with phosphate-buffered saline (PBS), the sections were blocked with an appropriate blocking solution to minimize nonspecific antibody binding. The sections were subsequently incubated overnight at 4°C with primary antibodies against IL11 and PDGFRα. After washing, appropriate fluorescence-conjugated secondary antibodies were applied and incubated in the dark. Cell nuclei were counterstained with DAPI. Following additional washing, the sections were mounted with an antifade mounting medium and examined using a fluorescence microscope. Representative images were acquired under identical imaging conditions for comparison between experimental groups. Colocalization of IL11 with PDGFRα-positive cells was assessed using merged fluorescence images.

2.17 Statistical Analysis

95% CI and P-values; enrichment analysis reports: P-values, p.adjust, and q-value; machine learning reports: ROC AUC and cross-validation results; expression analysis reports: mean and median values across groups, as well as P-values from Welch's t-test and the Wilcoxon/Mann-Whitney test; Immune infiltration analysis reports the correlation coefficient r and P-values; single-cell trend analysis reports cell-type-specific log2FC; virtual perturbation analysis reports perturbation significance; virtual screening and molecular docking reports predicted binding energies. All numerical values are derived from existing results tables or from re-calculations based on the expression matrix.

3. Results

3.1 Cadmium-Related Targets, Kidney Stone-Related Targets, and Transcriptomic Differences Jointly Identify

Candidate Molecules

The integration of toxicity targets yielded a total of 4,712 cadmium-related targets and 3,081 kidney stone-related targets. To further screen for candidate molecules at the transcriptomic level in diseased tissues, this study incorporated the GSE73680 dataset for differential expression analysis. This dataset included 33 cases of “Normal mucosa” and 29 cases of “Plaque mucosa.” Differential analysis revealed significant transcriptomic differences between Plaque mucosa and Normal mucosa, identifying a total of 730 differentially expressed genes, of which 301 were upregulated and 429 were downregulated (Fig. 1A). A heatmap of representative differentially expressed genes further demonstrated a certain degree of separation in the expression patterns of the two sample groups (Fig. 1B).

3.2 WGCNA and Intersection Analysis Prioritize Candidate Targets Containing IL11

A WGCNA analysis was performed using the GSE73680 expression matrix. The gene clustering tree and module color assignments showed that genes in the expression matrix could be divided into multiple modules based on their co-expression patterns (Fig. 1D). Module-trait correlation analysis revealed that the MEred module was positively correlated with “Plaque mucosa” ($r = 0.27$, $P = 0.03$) and negatively correlated with “Normal mucosa” ($r = -0.27$, $P = 0.03$) (Fig. 1C). Therefore, this study selected the MEred module as the module associated with the renal calculus-related tissue phenotype and extracted 205 genes from this module as WGCNA-related genes for subsequent multi-source intersection analysis.

After further integrating cadmium-related targets, kidney stone-related targets, differentially expressed genes, and MEred module genes, a total of 8 candidate targets were identified as intersections across the four categories, including IL11, TNFRSF11B, SPP1, AKR1B1, H19, PDZK1, SERPINA1, and CCK (Fig. 1E). These candidate targets link environmental cadmium exposure, the renal calculi disease context, tissue-specific differential expression, and WGCNA-related modules, suggesting that they may be involved in the tissue-level molecular changes associated with the risk of renal calculi due to urinary cadmium exposure. Among them, IL11 emerged repeatedly in differential expression analysis, intersection screening, and subsequent multilevel analysis; therefore, it was prioritized for inclusion in the exploration of mechanistic clues.

3.3 PPI Network and GO Enrichment Analysis of the 8 High-Confidence Candidate Targets

To further elucidate the interaction relationships and functional backgrounds of the 8 high-confidence candidate targets, this study constructed a PPI network and performed GO enrichment analysis. Since H19 is a lncRNA, it was excluded from the protein-protein interaction analysis; furthermore, PDZK1 did not form any major interaction links under the current network parameters. Therefore, the network diagram displays 6 candidate protein nodes, including SPP1, CCK, IL11, SERPINA1, TNFRSF11B, and

AKR1B1. The network comprises a total of 6 nodes and 10 edges, with nodes such as SPP1, CCK, IL11, and SERPINA1 exhibiting relatively high connectivity (Fig. 2A).

A GO enrichment analysis was further conducted on the 8 high-confidence candidate targets. Since H19 is a non-coding RNA and was not included in the GO enrichment analysis, the enrichment analysis was actually based on 7 annotated genes. The GO molecular function results showed that the candidate targets were primarily enriched in categories such as cytokine activity, aldose reductase (NADPH) activity, peptide hormone receptor binding, transporter activator activity, neuropeptide activity, hormone receptor binding, and oxidoreductase activity. Among these, “cytokine activity” was the most significantly enriched term, involving IL11, SPP1, and TNFRSF11B, with a GeneRatio of 3/7 and p.adjust = 0.0026, suggesting that this set of candidates may be associated with cytokine-related functions. No significantly enriched pathways were identified in the KEGG analysis after correction; therefore, this study does not consider the KEGG results as primary evidence for the underlying mechanism (Fig. 2B).

(A) PPI network diagram. H19 is a lncRNA; PDZK1 does not form major interactions, so the figure shows 6 candidate protein nodes and 10 interaction edges; node size represents the degree value. (B) GO molecular function enrichment analysis. Candidate targets were primarily enriched in functions such as cytokine activity, oxidoreductase activity, neuropeptide activity, and transporter activator activity; among these, cytokine activity involved IL11, SPP1, and TNFRSF11B. The KEGG analysis did not identify any statistically significant pathways after correction.

3.4 Machine Learning-Based Screening of IL-11 Using High-Confidence Candidate Targets

This study further conducted machine learning analyses focused on eight high-confidence candidate targets. Since this set simultaneously meets the criteria for cadmium-related targets, kidney stone disease targets, differentially expressed genes, and MEred module genes—and given the relatively limited number of features—it is particularly well-suited for prioritizing candidate core molecules within the sample size of GSE73680. Therefore, based on the expression matrices of these 8 high-confidence candidate targets in GSE73680, this study constructed LASSO, Random Forest, and SVM-RFE models, respectively.

LASSO regression determined the optimal λ parameter through cross-validation and screened candidate targets based on changes in coefficient paths (Fig. 3A, B). The random forest importance ranking showed that IL11 had the highest importance score among the candidate genes, followed by SPP1, H19, PDZK1, and TNFRSF11B (Fig. 3C). SVM-RFE (Recursive Feature Elimination) analysis revealed differences in model cross-validation accuracy across varying numbers of features, with the single-feature model performing relatively well, suggesting the presence of priority candidate molecules with strong discriminatory power within the candidate set (Fig. 3D).

Further evaluation of the models’ discriminatory ability using ROC curves showed that the RF, SVM, and LASSO models all possessed a certain degree of discriminatory ability, with AUC values of 0.757, 0.792,

and 0.778, respectively (Fig. 3E). After taking the intersection of the screening results from the three algorithms, it was found that IL11 was the only candidate gene selected by all three methods—LASSO, Random Forest, and SVM-RFE— (Fig. 3F). Therefore, this study identifies IL11 as a priority candidate core target and further explores its potential role through subsequent analyses of differential expression, immune relevance, GSEA, single-cell localization, and virtual perturbation analysis. It should be emphasized that machine learning was used in this study solely for candidate molecule ranking and not for clinical prediction or causal mechanism determination.

(A) LASSO regression cross-validation curve used to determine the optimal λ parameter for the model. (B) LASSO regression coefficient path plot showing changes in the coefficients of the 8 high-confidence candidate targets at different λ values. (C) Random Forest gene importance ranking plot, showing that IL11 has the highest importance score among the candidate genes. (D) SVM-RFE (Recursive Feature Elimination) validation plot, showing changes in the model's cross-validation accuracy at different numbers of features. (E) ROC curves for the three models—RF, SVM, and LASSO—with AUCs of 0.757, 0.792, and 0.778, respectively, indicating that the eight high-confidence candidate targets possess a certain ability to distinguish between "Plaque mucosa" and "Normal mucosa." (F) A Venn diagram showing the intersection of genes identified by the three algorithms—LASSO, Random Forest, and SVM-RFE—reveals that IL11 is a priority core candidate target common to all three algorithms.

3.5 Differences in IL-11 Expression, GSEA Analysis, and Immunological Correlation Analysis

Based on the results of multi-source intersection and machine learning screening, IL-11 was identified as a priority candidate core target. To further evaluate the expression profile of IL-11 in kidney stone-associated tissues, this study compared IL-11 expression levels between the "Normal mucosa" group and the "Plaque mucosa" group. The results showed that IL-11 was significantly elevated in the Plaque mucosa group, suggesting a potential association with tissue changes in Randall's plaque-related renal papillary lesions (Fig. 4A). KEGG-GSEA analysis using IL-11 expression levels as a phenotype revealed that genes associated with high IL-11 expression were primarily positively enriched in pathways such as lysosomes, focal adhesion, and cell adhesion molecules (Fig. 4B). Among these, the lysosome pathway suggests that IL-11-associated transcriptional features may be related to intracellular degradation, antigen processing, and inflammation-related intracellular homeostasis regulation; the focal adhesion and cell adhesion molecules pathways suggest that IL-11 may be associated with cell adhesion, extracellular matrix interactions, and local tissue microenvironment remodeling. Subsequently, this study analyzed the correlation between IL-11 expression and the level of immune cell infiltration. The results showed that IL-11 was significantly positively correlated with resting CD4 memory T cells, while it was negatively correlated with follicular helper T cells, eosinophils, memory B cells, and resting NK cells (Fig. 4C). These findings suggest that elevated IL-11 expression may be associated with changes in the composition of immune cells in kidney stone-associated tissues. Based on the combined findings of IL-11 expression differences, GSEA, and immune correlations, IL-11 may be involved in cell adhesion, tissue

microenvironment remodeling, and alterations in the composition of immune cells in kidney stone-associated tissues.

(A) Differences in IL-11 expression between the “Normal mucosa” and “Plaque mucosa” groups in the GSE73680 dataset. Box plots combined with a swarm plot show that IL-11 is significantly elevated in the “Plaque mucosa” group. (B) KEGG-GSEA analysis based on IL-11 expression levels. Genes highly co-expressed with IL11 were predominantly positively enriched in pathways such as KEGG_LYSOSOME, KEGG_FOCAL_ADHESION, and KEGG_CELL_ADHESION_MOLECULES_CAMS, suggesting that IL11 may be associated with lysosomal function, cell adhesion, extracellular matrix interactions, and local tissue microenvironment remodeling. (C) Correlation analysis between IL-11 expression and the level of immune cell infiltration. The x-axis represents the correlation coefficient, and the y-axis represents the immune cell type; bubble size indicates the absolute value of the correlation coefficient, and color indicates the P-value. The results show that IL-11 is significantly positively correlated with resting CD4 memory T cells, while it is negatively correlated with follicular helper T cells, eosinophils, memory B cells, and resting NK cells, suggesting that elevated IL-11 expression may be associated with changes in the composition of immune cells in kidney stone-associated tissues.

3.6 Single-Cell Localization and IL-11 Virtual Perturbation Analysis

To further clarify the cellular origin of IL-11 in kidney stone-associated tissues and its potential regulatory role, this study incorporated single-cell data from the GSE231569 renal papilla dataset for analysis. UMAP visualization results showed that cells in renal papillary tissue could be annotated as major cell populations including collecting duct principal cells, collecting duct intercalated cells, Loop of Henle TAL cells, papillary epithelial cells, endothelial cells, fibroblasts, pericytes, smooth muscle cells, macrophages/monocytes, and T cells (Fig. 5A).

Subsequently, this study analyzed the expression distribution of IL-11 across different cell types. Dot plot results showed that IL-11 exhibited relatively higher mean expression levels and a higher proportion of expressing cells in fibroblasts, whereas expression was lower or weaker in other cell types (Fig. 5B). These findings suggest that fibroblasts may be an important source of IL-11 expression in renal papillary tissue.

We further compared the trends in IL-11 expression across different cell types under the “Stone” and “Control” conditions. The results showed that IL-11 exhibited varying degrees of positive log2FC changes across multiple cell types, with relatively pronounced changes observed in cell populations such as macrophages/monocytes, fibroblasts, and Henle’s loop TAL cells (Fig. 5C). It should be noted that this heatmap is primarily used to illustrate the trends in IL-11 expression across different cell types and should not be interpreted on its own to imply that statistically significant differences were observed in every cell type.

Given the high expression of IL-11 in fibroblasts, this study further selected fibroblasts for scTenifoldKnk virtual perturbation analysis to predict potential IL-11-related transcriptional regulatory networks. This analysis constructed a gene regulatory network based on 600 downsampled fibroblasts and 1,200 candidate genes, and simulated changes in the downstream transcriptional network following IL-11 perturbation. The results showed that, following virtual IL-11 perturbation, a total of 22 genes reached a significant perturbation level with an FDR < 0.05, and 79 genes reached a perturbation level with P < 0.05 (Fig. 5D). In the volcano plot, genes such as ZBTB16, AP000331.1, ITGA2, FKBP5, INHBA, ANGPTL4, PLAT, CDH13, STC1, and PFKFB4 exhibited relatively pronounced changes, suggesting that IL-11 may be involved in the regulation of fibroblast-related transcriptional networks.

Further analysis of mechanism-related genes revealed that the IL11 virtual perturbation involves genes such as IL6ST, JAK1, STAT3, BRAF, TGFBR2, SMAD3, COL1A1, COL1A2, COL3A1, DCN, LUM, TIMP1, CCL2, SPP1, and HIF1A, among others. These genes are associated with the JAK/STAT-MAPK signaling pathway, the TGF-β/SMAD axis, extracellular matrix remodeling, and inflammatory damage responses. These findings suggest that IL-11 may be involved in the regulation of inflammation, matrix remodeling, and alterations in the local tissue microenvironment in renal papillary fibroblasts, thereby providing computational clues regarding candidate molecular mechanisms underlying the risk of kidney stones associated with cadmium exposure.

(A) UMAP visualization of single-cell data from renal papillae in GSE231569, showing annotation results for different cell types, including collecting duct principal cells, collecting duct intercalated cells, Loop of Henle TAL, papillary epithelial, endothelial cells, fibroblasts, pericytes, smooth muscle, macrophages/monocytes, and T cells, among other cell populations. (B) Dot plot showing the expression distribution of IL-11 across different cell types. Dot color indicates the average expression level, while dot size represents the proportion of cells expressing the gene. IL-11 is primarily highly expressed in fibroblasts. (C) Heatmap showing the expression trends of IL-11 across different cell types; colors represent the log2FC of Stone relative to Control. The results suggest that IL-11 exhibits varying degrees of upregulation across different cell types, with relatively pronounced changes observed in cell populations such as macrophages/monocytes, fibroblasts, and Loop of Henle TAL. (D) Volcano plot of the virtual perturbation analysis of IL11 in fibroblasts. The x-axis represents the perturbation score, and the y-axis represents −log10(P-value). Red points indicate significantly perturbed genes with FDR < 0.05, while blue points indicate perturbed genes with P < 0.05.

3.7 IL-11-Based Virtual Screening and Molecular Docking Analysis

To further explore potential small-molecule leads related to IL-11, this study conducted virtual screening and molecular docking analyses of 1,522 candidate small molecules with IL-11 as the target protein. Among the comprehensive screening results, ZINC000150338755 exhibited the lowest predicted binding energy; Based on the primary mechanism of IL11 and the interpretability of the candidate small molecules, this study selected Tetracycline/ZINC000084441937 as the mechanism-priority candidate

ligand for presentation. The results showed that the predicted binding energy between Tetracycline and IL11 was − 10.14 kcal/mol, suggesting that the two may form a relatively stable binding conformation. Combined with the aforementioned results from network toxicology, machine learning, single-cell localization, and IL-11 virtual perturbation, the low predicted binding energy between Tetracycline and IL-11 provides exploratory clues for subsequent protein-ligand binding validation and candidate small molecule research centered on IL-11. Molecular docking conformation analysis shows that tetracycline can fit into a localized binding pocket of the IL-11 protein and form predicted interactions near ARG-77. A local magnified view shows predicted hydrogen bond interactions between the ligand and amino acid residues near ARG-77, with bond lengths of approximately 2.4 Å and 2.6 Å, suggesting that these interactions may help maintain the stable conformation of the protein-ligand complex (Fig. 6).

A molecular docking analysis was performed using IL-11 as the receptor protein and tetracycline as the ligand. The overall conformation diagram shows that tetracycline can bind to a localized binding pocket on IL-11; the close-up view reveals predicted interactions between tetracycline and the region near ARG-77 on IL-11, with yellow dashed lines indicating predicted hydrogen bonds at distances of approximately 2.4 Å and 2.6 Å. The predicted binding energy between tetracycline and IL-11 is − 10.14 kcal/mol, suggesting that the two may form a relatively stable binding conformation.

3.8 Cadmium exposure aggravates CaOx crystal deposition and enhances fibroblast-associated IL11 expression in the kidney

To investigate the effect of cadmium (Cd) exposure on CaOx crystal-induced renal injury, mice were exposed to cadmium chloride ($CdCl_2$, 25 mg/L) in drinking water for 8 weeks, followed by glyoxylate (Gly) administration to induce renal CaOx crystal deposition. No detectable crystal deposition was observed in the kidneys of the saline group, whereas abundant birefringent CaOx crystals were present following Gly administration. Compared with the Gly group, the Gly + Cd group exhibited more extensive renal crystal deposition, indicating that Cd exposure aggravated Gly-induced CaOx crystal accumulation (Fig. 7A). Consistently, PAS staining showed preserved renal tubular morphology in the saline group, whereas Gly treatment resulted in evident tubular dilation and epithelial injury. These histopathological alterations were further aggravated in the Gly + Cd group (Fig. 7B). Immunofluorescence staining further showed that PDGFRα and IL11 signals were weak in the saline group but markedly increased following Gly treatment. Compared with the Gly group, the Gly + Cd group exhibited further enhancement of both PDGFRα and IL11 fluorescence signals, accompanied by increased spatial colocalization of IL11 with PDGFRα-positive fibroblasts (Fig. 7C). Together, these findings demonstrate that Cd exposure increases renal CaOx crystal deposition and aggravates tubular injury, accompanied by enhanced IL11 expression in PDGFRα-positive fibroblasts.

(A) Representative images of renal calcium oxalate (CaOx) crystal deposition in the Saline, glyoxylate (Gly), and Gly plus cadmium (Gly + Cd) groups. Kidney sections were examined under polarized light microscopy, and the lower panels show magnified views of the indicated regions. (B) Representative

periodic acid–Schiff (PAS)-stained kidney sections showing renal tubular histopathological alterations among the three groups. (C) Representative immunofluorescence images showing the expression and distribution of PDGFRα (red) and IL11 (green) in kidney tissues from the Saline, Gly, and Gly + Cd groups. Nuclei were counterstained with DAPI (blue). Merged images show the spatial colocalization of IL11 with PDGFRα-positive fibroblasts.

4. Discussion

This study integrates multidimensional bioinformatics analysis strategies to screen for potential biomarkers of kidney stones and systematically analyzes the molecular associations between environmental cadmium exposure and the development of kidney stones. It conducts a multilevel functional analysis of the core candidate gene IL11 to explore the potential molecular mechanisms by which cadmium exposure, mediated by this gene, contributes to the development and progression of kidney stones, thereby providing a theoretical framework and preliminary insights for subsequent in vivo and in vitro functional validation and the identification of pathogenic pathways.

Cadmium accumulates in the kidneys, and long-term exposure may affect renal tubular function, oxidative stress, inflammatory responses, and the local tissue microenvironment[33, 34]. The formation of kidney stones is not merely a process of crystal deposition but involves multiple steps, including damage to the renal tubular epithelium, crystal adhesion, local inflammation, the involvement of immune cells, and remodeling of the renal papillae. Therefore, integrating cadmium exposure with transcriptional changes in kidney stone-associated tissues can help propose more verifiable candidate mechanism hypotheses. In this study, by intersecting cadmium-related targets, kidney stone disease targets, DEGs, and genes from the MEred module, we identified eight high-confidence candidate targets from the intersection of these four categories. Previous studies have also analyzed Randall's plaque-associated differentially expressed genes, WGCNA modules, PPI networks, and immune infiltration based on GSE73680, suggesting that this dataset can be used to explore the molecular characteristics of renal papillary lesions associated with kidney stones[3, 35]. This screening strategy requires candidate molecules to simultaneously link environmental cadmium exposure, the kidney stone disease context, tissue differential expression, and WGCNA-related modules; therefore, it is more integrative than relying solely on database targets or differentially expressed genes. Subsequent PPI, GO, and machine learning analyses were all centered on these 8 candidate targets, focusing the main research thread on "8 high-confidence candidate targets—IL-11 prioritization—IL-11-related cells and signaling pathways."

The PPI network of the 8 candidate targets shows that SPP1, CCK, IL11, SERPINA1, TNFRSF11B, and AKR1B1 form the primary interaction network. Since H19 is a lncRNA and PDZK1 does not form any major interactions, the PPI diagram actually displays 6 protein nodes. IL11 exhibits high connectivity within the network but is not the node with the highest connectivity; therefore, it cannot be defined as the primary core target based solely on PPI analysis. GO molecular function enrichment analysis indicates that the candidate targets are primarily associated with cytokine activity, including IL11, SPP1, and TNFRSF11B, suggesting that this set of candidates is related to cytokine-associated functions. No

corrected, significant KEGG pathways were identified for the eight candidate targets, which may be related to the small number of genes and limited pathway coverage; therefore, this study does not consider KEGG as the primary basis for identifying mechanisms. Machine learning results further support IL11's status as a priority candidate. Among the eight high-confidence candidate targets, IL11 was the only gene selected by all three algorithms—LASSO, Random Forest, and SVM-RFE—suggesting that this gene plays a significant role in the relationship between cadmium exposure and the development of kidney stones.

Further analysis of IL-11 revealed that IL-11 was significantly upregulated in plaque mucosa and correlated with the proportion of certain immune cell infiltrates. GSEA results showed that genes associated with high IL-11 expression were primarily positively enriched in pathways such as lysosomes, focal adhesions, and cell adhesion molecules, suggesting a potential association with intracellular degradation, antigen processing, cell adhesion, extracellular matrix interactions, and local tissue microenvironment remodeling. These findings are associated with pathological processes in kidney stone formation, such as epithelial damage, crystal adhesion, local inflammation, and tissue remodeling, providing a hypothetical basis for further exploration of the "IL-11–cell adhesion/matrix remodeling–changes in the renal papillary microenvironment" pathway.

Single-cell analysis provided further clues regarding the cellular localization of IL-11. In the GSE231569 renal papilla single-cell dataset, IL-11 was primarily highly expressed in fibroblasts, suggesting that fibroblasts may be an important source of IL-11 expression in renal papilla tissue associated with kidney stones. Cell-type-specific log2FC heatmaps show that IL-11 exhibits varying degrees of positive upregulation across multiple cell populations; however, these results are primarily intended to describe trends and should not be interpreted to mean that IL-11 is significantly upregulated in all cell types. Based on the expression profile of IL-11 in fibroblasts, this study further conducted a scTenifoldKnk virtual perturbation analysis. The results suggest that IL-11 perturbation may affect gene networks related to JAK/STAT-MAPK signaling, the TGF-β/SMAD axis, extracellular matrix remodeling, and inflammatory damage. There is a body of literature linking IL-11 to renal fibrosis, fibroblast activation, and tissue remodeling[36–40]. Although these studies did not directly address kidney stones, they suggest that IL-11 may be involved in cellular state changes and matrix remodeling in pathological renal conditions. In this study, the elevated expression of IL-11, its localization in fibroblasts, the GSEA pathway clues, and the results of virtual perturbation analysis show a degree of consistency, collectively supporting the hypothesis that IL-11 may be involved in changes to the renal papillary microenvironment associated with kidney stones.

The results of the virtual screening provide an exploratory supplement to potential IL-11-related small-molecule leads. In this study, Tetracycline/ZINC000084441937 was selected as a mechanism-priority candidate ligand for presentation. Its predicted binding energy with IL-11 is − 10.14 kcal/mol, and it forms predicted hydrogen bond interactions near ARG-77, suggesting that Tetracycline may have the potential to bind to IL-11.

This study still has certain limitations. First, the cadmium-related targets and kidney stone-related targets were derived from databases, and there are differences in the level of evidence and coverage among different databases. Second, the GSE73680 dataset has a limited sample size, and it primarily reflects local transcriptional differences between Randall's plaque-associated renal papillary tissue and non-plaque renal papillary tissue; it cannot be simply equated with a patient-level comparison of "kidney stone group vs. healthy control group." Therefore, the machine learning results can only be used for candidate ranking. Third, single-cell results are primarily used for cell localization and trend description; both scTenifoldKnk and molecular docking are computational predictions and cannot replace actual functional experiments.

In summary, this study identified eight high-confidence candidate targets involved in the pathological processes of cadmium exposure and kidney stones. Through multi-level bioinformatics analyses—including intersection analysis, protein interaction networks, and machine learning-based prioritization—IL11 was further identified as the core research target. The findings related to IL-11 primarily point to cellular adhesion, lysosomal function, immune-related changes, and extracellular matrix remodeling processes in the context of fibroblasts. These results provide potential avenues for understanding the underlying tissue microenvironment mechanisms underlying the risk of kidney stones associated with cadmium exposure and lay the groundwork for subsequent functional validation studies focusing on the IL-11–fibroblasts–matrix remodeling axis.

Declarations

Conflict of interests

The authors declare no conflict of interest.

Ethics approval and consent to participate

All animal experiments in this study were approved by the Institutional Animal Care and Use Committee of Anhui Medical University (Approval No. LLSC20251958).

Consent to Participate

Not applicable, as this study does not involve human participants.

Consent for publication

Not applicable.

Funding Declaration

This research did not receive funding.

Author Contribution

Conceptualization, ZZ and ZH; methodology, ZZ and XZ; software, LZ and ZC; formal analysis,LZand ZC; data curation, ZZ; writing—original draft preparation, ZZand XZ.; writing—review and editing, ZZ and ZH. All authors have read and agreed to the published version of the manuscript.

Data Availability

All data supporting the findings of this study are available within the paper .

References

1. Jiang H, Jiang Y, Ji J, Chou A, Mao W, Chen M, Zheng Y. Upregulation of COL3A1 is a Key Driver in Kidney Stone Formation. Front Biosci (Landmark Ed). 2026;31(5):51232. 10.31083/FBL51232.
2. Wang N, Zhou B, Xu C, Hao B. Identification of Biomarker in Kidney Stone Disease by Integrating Transcriptomics and Olink Proteomics: A Case-Control Study. J Inflamm Res. 2026;19:605025. 10.2147/JIR.S605025.
3. Hong SY, Xia QD, Xu JZ, Liu CQ, Sun JX, Xun Y, Wang SG. Identification of the pivotal role of SPP1 in kidney stone disease based on multiple bioinformatics analysis. BMC Med Genomics. 2022;15(1):7. 10.1186/s12920-022-01157-4.
4. Wang M, Chen Z, Song W, Hong D, Huang L, Li Y. A review on Cadmium Exposure in the Population and Intervention Strategies Against Cadmium Toxicity. Bull Environ Contam Toxicol. 2021;106(1):65–74. 10.1007/s00128-020-03088-1.
5. Adams SV, Newcomb PA, Shafer MM, Atkinson C, Bowles EJ, Newton KM, Lampe JW. Sources of cadmium exposure among healthy premenopausal women. Sci Total Environ. 2011;409(9):1632–7. 10.1016/j.scitotenv.2011.01.037.
6. Doccioli C, Sera F, Francavilla A, Cupisti A, Biggeri A. Association of cadmium environmental exposure with chronic kidney disease: A systematic review and meta-analysis. Sci Total Environ. 2024;906:167165. 10.1016/j.scitotenv.2023.167165.
7. Kaewnate Y, Niyomtam S, Tangvarasittichai O, Meemark S, Pingmuangkaew P, Tangvarasittichai S. Association of elevated urinary cadmium with urinary stone, hypercalciuria and renal tubular dysfunction in the population of cadmium-contaminated area. Bull Environ Contam Toxicol. 2012;89(6):1120–4. 10.1007/s00128-012-0856-8.

8. Satarug S, Garrett SH, Sens MA, Sens DA. Cadmium, environmental exposure, and health outcomes. Environ Health Perspect. 2010;118(2):182–90. 10.1289/ehp.0901234.
9. Davis AP, Wiegers TC, Sciaky D, Barkalow F, Strong M, Wyatt B, Wiegers J, McMorran R, Abrar S, Mattingly CJ. Comparative Toxicogenomics Database's 20th anniversary: update 2025. Nucleic Acids Res. 2025;53(D1):D1328–34. 10.1093/nar/gkae883.
10. Safran M, Dalah I, Alexander J, Rosen N, Iny Stein T, Shmoish M, Nativ N, Bahir I, Doniger T, Krug H, Sirota-Madi A, Olender T, Golan Y, Stelzer G, Harel A, Lancet D. GeneCards Version 3: the human gene integrator. Database (Oxford). 2010;2010:baq020. 10.1093/database/baq020.
11. Amberger JS, Bocchini CA, Schiettecatte F, Scott AF, Hamosh A. OMIM.org: Online Mendelian Inheritance in Man (OMIM®), an online catalog of human genes and genetic disorders. Nucleic Acids Res. 2015;43(Database issue):D789–98. 10.1093/nar/gku1205.
12. Taguchi K, Hamamoto S, Okada A, Unno R, Kamisawa H, Naiki T, Ando R, Mizuno K, Kawai N, Tozawa K, Kohri K, Yasui T. Genome-Wide Gene Expression Profiling of Randall's Plaques in Calcium Oxalate Stone Formers. J Am Soc Nephrol. 2017;28(1):333–47. 10.1681/ASN.2015111271.
13. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008;9:559. 10.1186/1471-2105-9-559.
14. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet. 2000;25(1):25–9. 10.1038/75556.
15. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 2000;28(1):27–30. 10.1093/nar/28.1.27.
16. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS. 2012;16(5):284–7. 10.1089/omi.2011.0118.
17. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, Feng T, Zhou L, Tang W, Zhan L, Fu X, Liu S, Bo X, Yu G. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innov (Camb). 2021;2(3):100141. 10.1016/j.xinn.2021.100141.
18. Hao PY, Chiang JH, Chen YD. Possibilistic classification by support vector networks. Neural Netw. 2022;149:40–56. 10.1016/j.neunet.2022.02.007.
19. Daneshvar A, Golalizadeh M. Quantile regression shrinkage and selection via the Lqsso. J Biopharm Stat. 2024;34(3):297–322. 10.1080/10543406.2023.2198593.
20. Becker T, Rousseau AJ, Geubbelmans M, Burzykowski T, Valkenborg D. Decision trees and random forests. Am J Orthod Dentofac Orthop. 2023;164(6):894–7. 10.1016/j.ajodo.2023.09.011.
21. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, Müller M. pROC: an open-source package for R and S + to analyze and compare ROC curves. BMC Bioinformatics. 2011;12:77. 10.1186/1471-2105-12-77.
22. Rong Y, Zhao SD, Zheng X, Li Y. Kernel Cox partially linear regression: Building predictive models for cancer patients' survival. Stat Med. 2024;43(1):1–15. 10.1002/sim.9938.

23. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M, Alizadeh AA. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods. 2015;12(5):453–7. 10.1038/nmeth.3337.
24. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005;102(43):15545–50. 10.1073/pnas.0506580102.
25. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst. 2015;1(6):417–25. 10.1016/j.cels.2015.12.004.
26. Canela VH, Bowen WS, Ferreira RM, Syed F, Lingeman JE, Sabo AR, Barwinska D, Winfree S, Lake BB, Cheng YH, Gaut JP, Ferkowicz M, LaFavers KA, Zhang K, Coe FL, Worcester E, Jain S, Eadon MT, Williams JC Jr, El-Achkar TM. A spatially anchored transcriptomic atlas of the human kidney papilla identifies significant immune injury in patients with stone disease. Nat Commun. 2023;14(1):4140. 10.1038/s41467-023-38975-8.
27. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat Biotechnol. 2018;36(5):411–20. 10.1038/nbt.4096.
28. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM 3rd, Hao Y, Stoeckius M, Smibert P, Satija R. Comprehensive Integration of Single-Cell Data. Cell. 2019;177(7):1888–e190221. 10.1016/j.cell.2019.05.031.
29. Hao Y, Hao S, Andersen-Nissen E, Mauck WM 3rd, Zheng S, Butler A, Lee MJ, Wilk AJ, Darby C, Zager M, Hoffman P, Stoeckius M, Papalexi E, Mimitou EP, Jain J, Srivastava A, Stuart T, Fleming LM, Yeung B, Rogers AJ, McElrath JM, Blish CA, Gottardo R, Smibert P, Satija R. Integrated analysis of multimodal single-cell data. Cell. 2021;184(13):3573–e358729. 10.1016/j.cell.2021.04.048.
30. Osorio D, Zhong Y, Li G, Xu Q, Yang Y, Tian Y, Chapkin RS, Huang JZ, Cai JJ. scTenifoldKnk: An efficient virtual knockout tool for gene function predictions via single-cell gene regulatory network perturbation. Patterns (N Y). 2022;3(3):100434. 10.1016/j.patter.2022.100434
31. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem. 2010;31(2):455–61. 10.1002/jcc.21334.
32. Eberhardt J, Santos-Martins D, Tillack AF, Forli S, AutoDock. Vina 1.2.0: New Docking Methods, Expanded Force Field, and Python Bindings. J Chem Inf Model. 2021;61(8):3891–8. 10.1021/acs.jcim.1c00203.
33. Johri N, Jacquillet G, Unwin R. Heavy metal poisoning: the effects of cadmium on the kidney. Biometals. 2010;23(5):783–92. 10.1007/s10534-010-9328-y.
34. Yan LJ, Allen DC. Cadmium-Induced Kidney Injury: Oxidative Damage as a Unifying Mechanism. Biomolecules. 2021;11(11):1575. 10.3390/biom11111575.

35. Yu L, Li G, Jin S, Su J, Li S. Identification of the core genes in Randall's plaque of kidney stone and immune infiltration with WGCNA network. Front Genet. 2023;14:1048919. 10.3389/fgene.2023.1048919.
36. Widjaja AA, Viswanathan S, Shekeran SG, Adami E, Lim WW, Chothani S, Tan J, Goh JWT, Chen HM, Lim SY, Boustany-Kari CM, Hawkins J, Petretto E, Hübner N, Schafer S, Coffman TM, Cook SA. Targeting endogenous kidney regeneration using anti-IL11 therapy in acute and chronic models of kidney disease. Nat Commun. 2022;13(1):7497. 10.1038/s41467-022-35306-1.
37. Widjaja AA, Shekeran SG, Adami E, Ting JGW, Tan J, Viswanathan S, Lim SY, Tan PH, Hübner N, Coffman T, Cook SA. A Neutralizing IL-11 Antibody Improves Renal Function and Increases Lifespan in a Mouse Model of Alport Syndrome. J Am Soc Nephrol. 2022;33(4):718–30. 10.1681/ASN.2021040577.
38. Widjaja AA, Cook SA. The Pathobiology of IL-11 in Kidney Disease: From Epithelial Cell to Fibroblast and Back Again. Am J Pathol. 2023;193(12):1910–3. 10.1016/j.ajpath.2023.09.011.
39. Li Y, Luo C, Zeng Y, Zheng Z, Tao D, Liu Q, Hong Y, Wang S, Long H, Xu Z. Renal Fibrosis Is Alleviated through Targeted Inhibition of IL-11-Induced Renal Tubular Epithelial-to-Mesenchymal Transition. Am J Pathol. 2023;193(12):1936–52. 10.1016/j.ajpath.2023.07.005.
40. Pang N, Zhao N, An C, Li K, Li P, Wang N, Li J, Cheng X, Zheng N, Guo D, Xiong X. Development of a Long-Acting Interleukin-11 Antagonist for the Treatment of Renal Fibrosis. J Med Chem. 2025;68(8):8429–38. 10.1021/acs.jmedchem.4c03185.

Figures

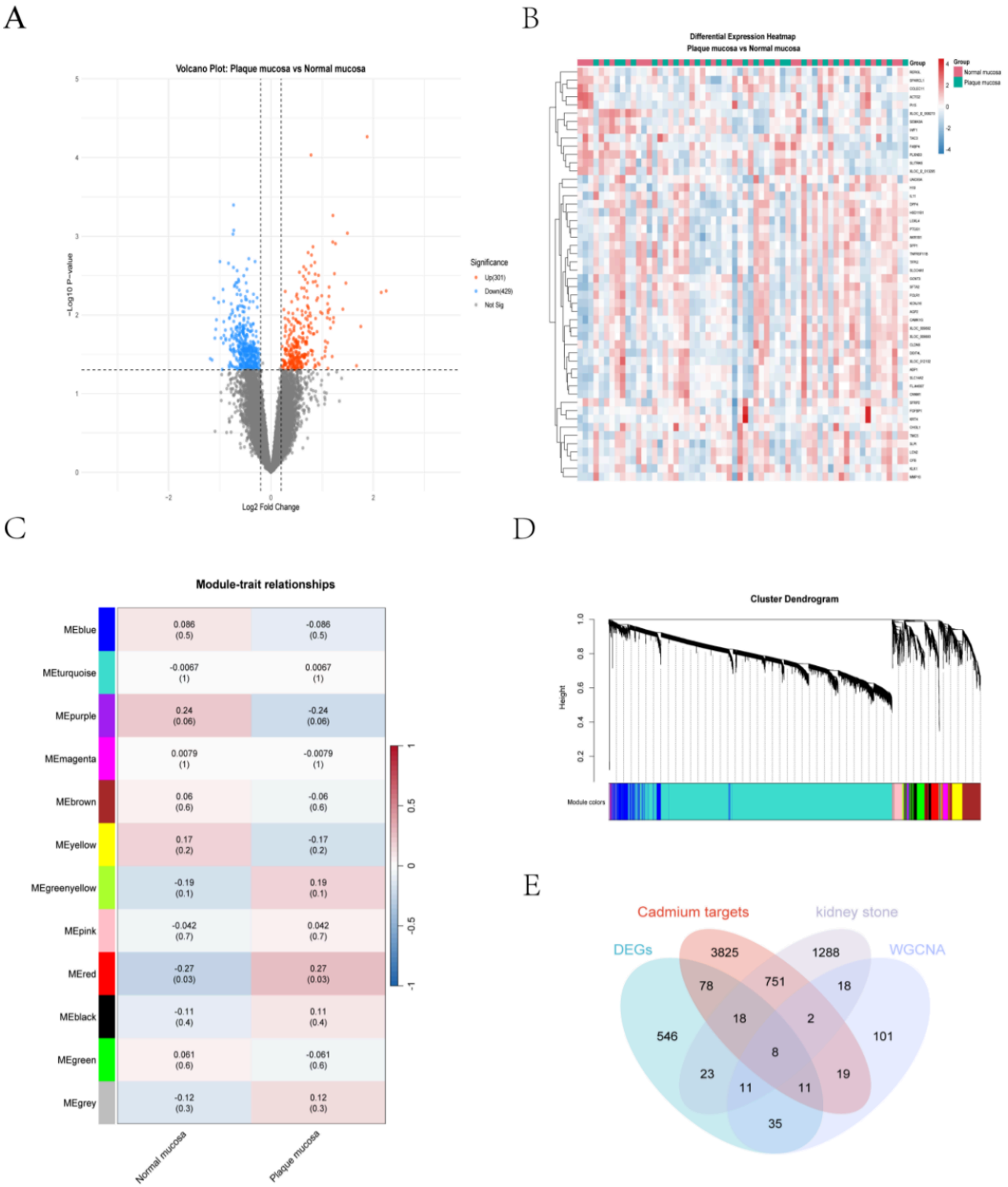


Figure 1

Transcriptomic differences, WGCNA, and screening of 8 high-confidence candidate targets. (A) Volcano plot of differentially expressed genes between “Plaque mucosa” and “Normal mucosa” in the GSE73680 dataset; red indicates upregulated genes, blue indicates downregulated genes, and gray indicates genes with no significant difference. (B) Heatmap of representative differentially expressed genes, showing expression differences between Plaque mucosa and Normal mucosa. The color bar at

the top indicates sample grouping, and the heatmap colors represent normalized expression levels. (C) WGCNA module-trait correlation heatmap, showing the correlations between different co-expression modules and tissue phenotypes. Among these, the MEred module is positively correlated with plaque mucosa (r = 0.27, P = 0.03) and negatively correlated with normal mucosa (r = −0.27, P = 0.03). (D) WGCNA gene clustering tree and module color assignments, showing genes grouped into different modules based on their co-expression patterns. (E) A Venn diagram showing the intersections of four categories—cadmium-related targets, kidney stone disease targets, differentially expressed genes, and genes in the MEred module—yielding a total of 8 overlapping candidate targets. The “WGCNA gene set” in the Venn diagram refers to the 205 genes in the MEred module.

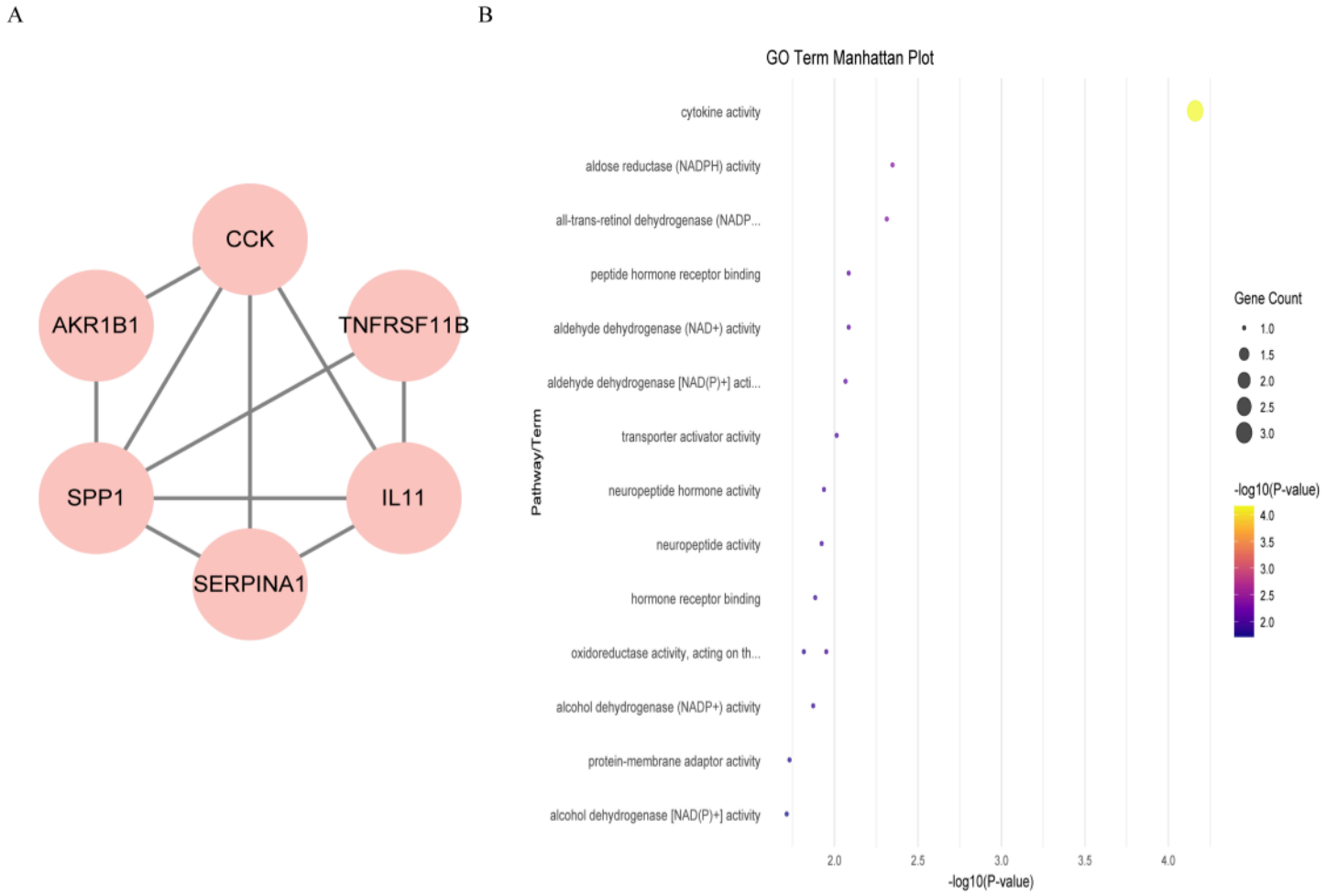


Figure 2

PPI network and GO enrichment analysis of 8 high-confidence candidate targets.

(A) PPI network diagram. H19 is a lncRNA; PDZK1 does not form major interactions, so the figure shows 6 candidate protein nodes and 10 interaction edges; node size represents the degree value. (B) GO molecular function enrichment analysis. Candidate targets were primarily enriched in functions such as cytokine activity, oxidoreductase activity, neuropeptide activity, and transporter activator activity; among

these, cytokine activity involved IL11, SPP1, and TNFRSF11B. The KEGG analysis did not identify any statistically significant pathways after correction.

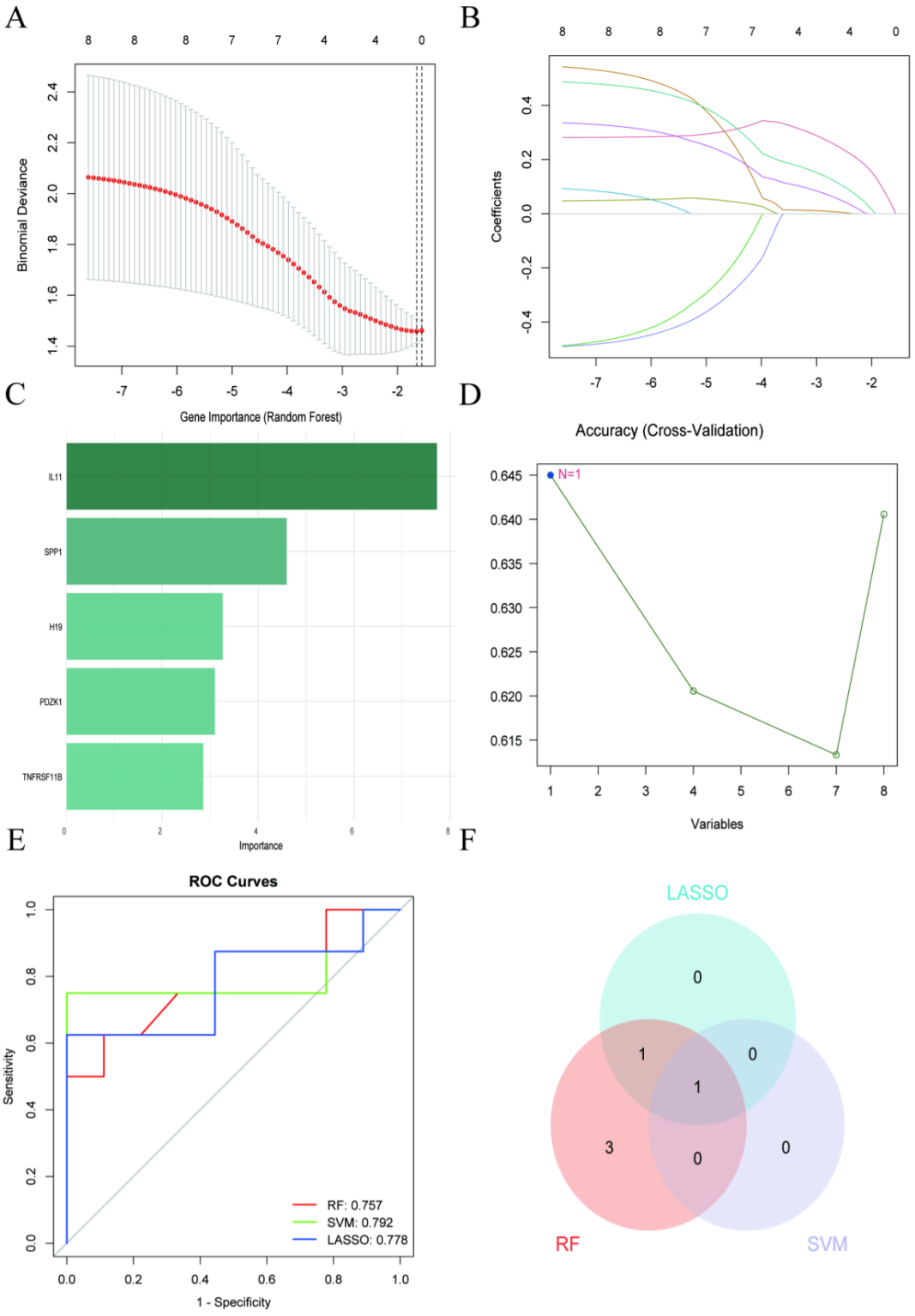


Figure 3

Machine learning-based prioritization analysis of 8 high-confidence candidate targets.

(A) LASSO regression cross-validation curve used to determine the optimal λ parameter for the model. (B) LASSO regression coefficient path plot showing changes in the coefficients of the 8 high-confidence candidate targets at different λ values. (C) Random Forest gene importance ranking plot, showing that IL11 has the highest importance score among the candidate genes. (D) SVM-RFE (Recursive Feature Elimination) validation plot, showing changes in the model's cross-validation accuracy at different numbers of features. (E) ROC curves for the three models—RF, SVM, and LASSO—with AUCs of 0.757, 0.792, and 0.778, respectively, indicating that the eight high-confidence candidate targets possess a certain ability to distinguish between "Plaque mucosa" and "Normal mucosa." (F) A Venn diagram showing the intersection of genes identified by the three algorithms—LASSO, Random Forest, and SVM-RFE—reveals that IL11 is a priority core candidate target common to all three algorithms.

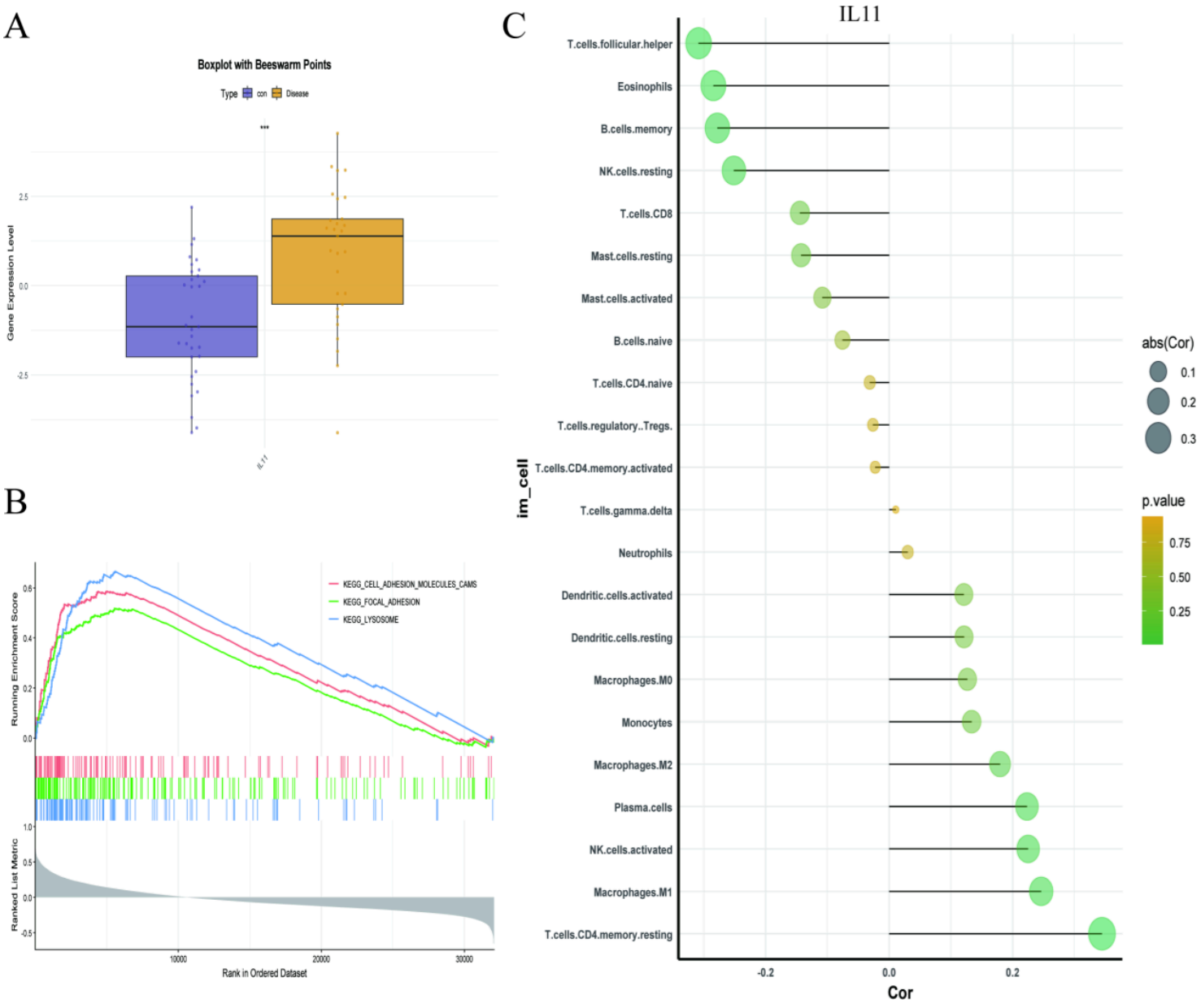


Figure 4

Differences in IL-11 expression, GSEA analysis, and correlations with immune cells.

(A) Differences in IL-11 expression between the “Normal mucosa” and “Plaque mucosa” groups in the GSE73680 dataset. Box plots combined with a swarm plot show that IL-11 is significantly elevated in the “Plaque mucosa” group. (B) KEGG-GSEA analysis based on IL-11 expression levels. Genes highly co-expressed with IL11 were predominantly positively enriched in pathways such as KEGG_LYSOSOME, KEGG_FOCAL_ADHESION, and KEGG_CELL_ADHESION_MOLECULES_CAMS, suggesting that IL11 may be associated with lysosomal function, cell adhesion, extracellular matrix interactions, and local tissue microenvironment remodeling. (C) Correlation analysis between IL-11 expression and the level of immune cell infiltration. The x-axis represents the correlation coefficient, and the y-axis represents the immune cell type; bubble size indicates the absolute value of the correlation coefficient, and color indicates the P-value. The results show that IL-11 is significantly positively correlated with resting CD4 memory T cells, while it is negatively correlated with follicular helper T cells, eosinophils, memory B cells, and resting NK cells, suggesting that elevated IL-11 expression may be associated with changes in the composition of immune cells in kidney stone-associated tissues.

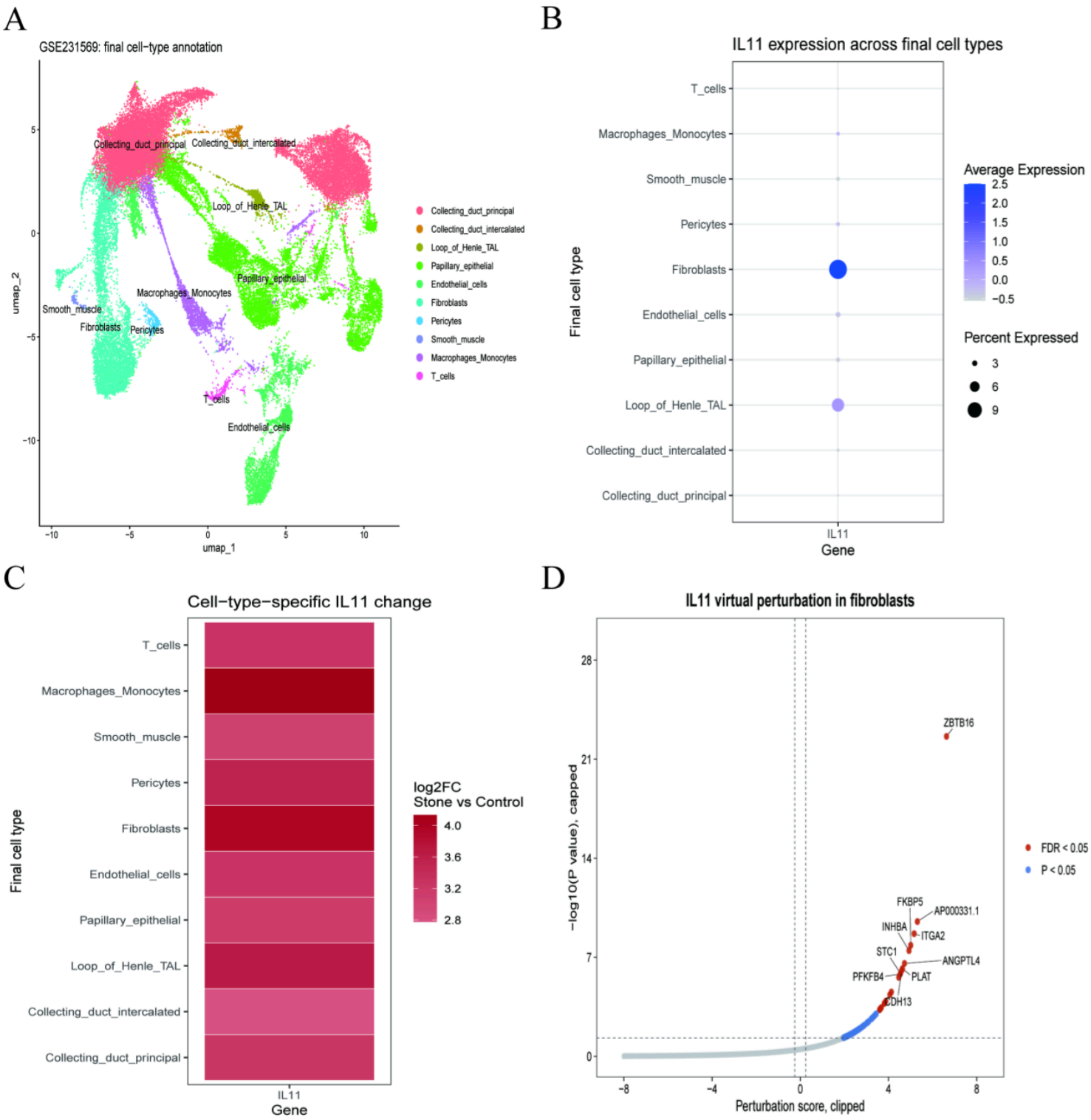


Figure 5

Single-cell localization and IL-11 virtual perturbation analysis.

(A) UMAP visualization of single-cell data from renal papillae in GSE231569, showing annotation results for different cell types, including collecting duct principal cells, collecting duct intercalated cells, Loop of Henle TAL, papillary epithelial, endothelial cells, fibroblasts, pericytes, smooth muscle, macrophages/monocytes, and T cells, among other cell populations. (B) Dot plot showing the

expression distribution of IL-11 across different cell types. Dot color indicates the average expression level, while dot size represents the proportion of cells expressing the gene. IL-11 is primarily highly expressed in fibroblasts. (C) Heatmap showing the expression trends of IL-11 across different cell types; colors represent the log2FC of Stone relative to Control. The results suggest that IL-11 exhibits varying degrees of upregulation across different cell types, with relatively pronounced changes observed in cell populations such as macrophages/monocytes, fibroblasts, and Loop of Henle TAL. (D) Volcano plot of the virtual perturbation analysis of IL11 in fibroblasts. The x-axis represents the perturbation score, and the y-axis represents −log10(P-value). Red points indicate significantly perturbed genes with FDR < 0.05, while blue points indicate perturbed genes with P < 0.05.

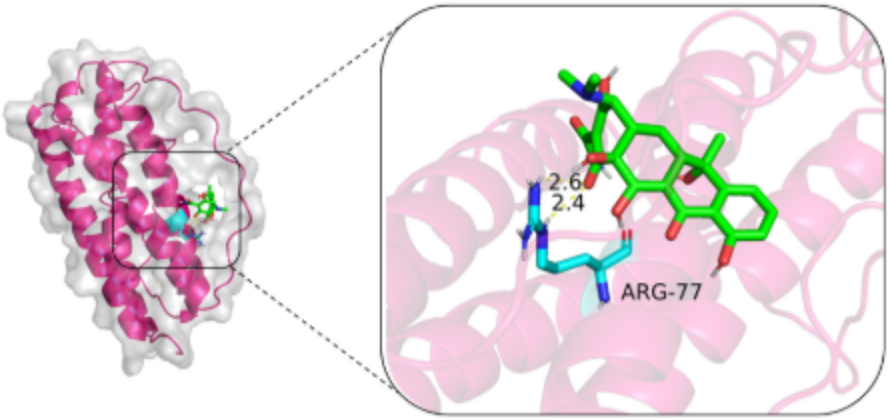


Figure 6

Molecular docking conformation of tetracycline with IL-11.

A molecular docking analysis was performed using IL-11 as the receptor protein and tetracycline as the ligand. The overall conformation diagram shows that tetracycline can bind to a localized binding pocket on IL-11; the close-up view reveals predicted interactions between tetracycline and the region near ARG-77 on IL-11, with yellow dashed lines indicating predicted hydrogen bonds at distances of approximately 2.4 Å and 2.6 Å. The predicted binding energy between tetracycline and IL-11 is −10.14 kcal/mol, suggesting that the two may form a relatively stable binding conformation.

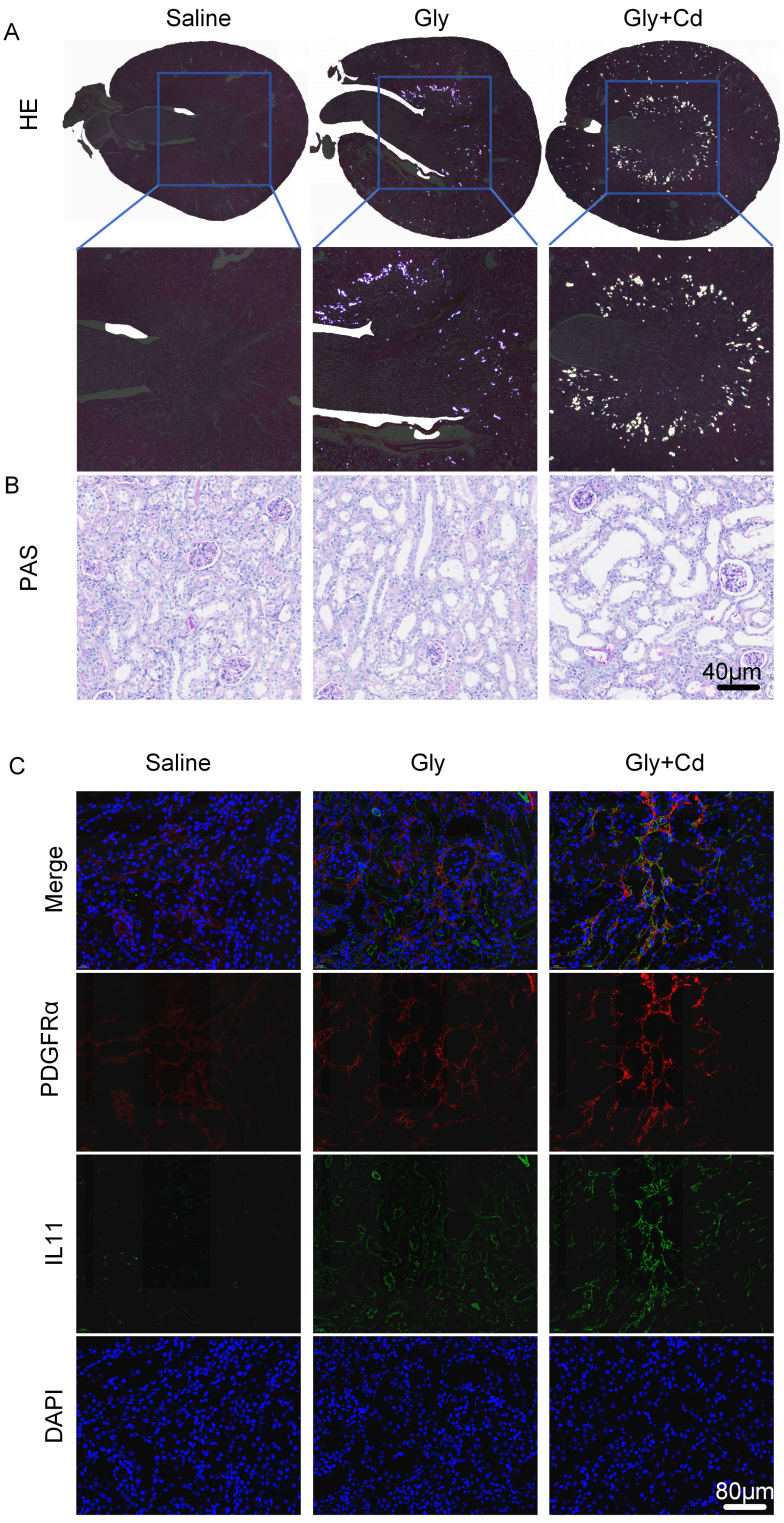


Figure 7

Cadmium exposure aggravates renal CaOx crystal deposition and is associated with enhanced IL11 expression in PDGFRα-positive fibroblasts.

(A) Representative images of renal calcium oxalate (CaOx) crystal deposition in the Saline, glyoxylate (Gly), and Gly plus cadmium (Gly+Cd) groups. Kidney sections were examined under polarized light

microscopy, and the lower panels show magnified views of the indicated regions. (B) Representative periodic acid–Schiff (PAS)-stained kidney sections showing renal tubular histopathological alterations among the three groups. (C) Representative immunofluorescence images showing the expression and distribution of PDGFRα (red) and IL11 (green) in kidney tissues from the Saline, Gly, and Gly+Cd groups. Nuclei were counterstained with DAPI (blue). Merged images show the spatial colocalization of IL11 with PDGFRα-positive fibroblasts.