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.

Network-Guided Systems Biology Prioritizes Regulatory Modulators, Potential Drug Targets, and Multi-Target Natural Candidate Compounds for Diabetic Nephropathy

Abhay Raj Kori
University of Allahabad
Mohammad Sahil Khan
University of Allahabad
Neha Singh
University of Allahabad
Pramod Katara
pmkatara@gmail.com

University of Allahabad

Research Article

Keywords: Diabetic Nephropathy, Molecular Network, Regulatory Modulators, Drug Targets, Therapeutic Compounds, Differentially Expressed Genes, Network Pharmacology

Posted Date: October 6th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10788745/v1

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

Additional Declarations: No competing interests reported.

Network-Guided Systems Biology Prioritizes Regulatory Modulators, Potential Drug Targets, and Multi-Target Natural Candidate Compounds for Diabetic Nephropathy

Abhay Raj Kori, Mohammad Sahil Khan, Neha Singh, Pramod Katara*

Author Affiliation

Abhay Raj Kori,
koriabhayraj063@gmail.com
Centre of Bioinformatics, University of Allahabad-
Prayagraj (UP), INDIA

Mohammad Sahil Khan,
msahilkhan0129@gmail.com
Centre of Bioinformatics, University of Allahabad-
Prayagraj (UP), INDIA

Neha Singh
ns028066@gmail.com
Centre of Bioinformatics,University of Allahabad
Prayagraj-(UP) INDIA

Pramod Katara
(orcid:0000-0002-3444-2822)
pmkatara@gmail.com, Pkatara@allduniv.ac.in
Centre of Bioinformatics, University of Allahabad-
Prayagraj (UP), INDIA

Corresponding Author*: Dr. Pramod Katara (pmkatara@gmail.com, Pkatara@allduniv.ac.in)

Abstract:

Diabetic nephropathy (DN) is a leading cause of end-stage renal disease, driven by metabolic dysregulation, chronic inflammation, and fibrotic remodeling. Elucidating the molecular networks underlying DN progression is essential for prioritizing therapeutic strategies. In this study, we employed an integrated *in silico* approach combining transcriptomic analysis, molecular network reconstruction, and structure-based virtual screening to characterize candidate disease-associated gene modules and prioritize multi-target natural compounds. RNA sequencing data from human kidney biopsies were analyzed. To prioritize hyperactive pathological cascades for pharmacological inhibition, downstream network clustering and virtual screening focused exclusively on upregulated differentially expressed genes (DEGs). Intersection with curated metabolic drug targets identified disease-associated genes. Protein–protein interaction and regulatory network analyses revealed subnetworks enriched for immune activation, inflammatory signaling, and extracellular matrix organization, highlighting highly connected nodes including *CD4, CCR7, CSF1R, FBN1, ITGA4, LCK, SYK,* and *TLR6.* Candidate regulatory elements, including transcription factors *(NFKB1, RELA, SP1, STAT3*) and miRNAs, were identified as potential modulators of these transcriptional programs. To bridge these network findings to therapeutic discovery, natural compounds from the LOTUS database were screened against 29 prioritized DN-associated protein targets using molecular docking. Network analysis prioritized ergostane-type steroids as candidate multi-target antagonists exhibiting favorable theoretical binding affinities. Collectively, these findings delineate key molecular networks in DN and prioritize mechanistically grounded therapeutic candidates, generating robust, data-driven hypotheses for future experimental validation.

Keywords: Diabetic Nephropathy, Molecular Network, Regulatory Modulators, Drug Targets, Therapeutic Compounds, Differentially Expressed Genes, Network Pharmacology

1. INTRODUCTION

Metabolic disorders, particularly diabetes mellitus, constitute a major global health challenge, affecting hundreds of millions of individuals and contributing substantially to morbidity and mortality worldwide (World Health Organization, 2024). Among the diverse complications associated with diabetes, microvascular disorders are particularly debilitating, with Diabetic Nephropathy (DN) emerging as a leading cause of end-stage renal disease and affecting nearly 40% of individuals with diabetes (Reutens & Atkins, 2011). Estimates from the Global Burden of Disease (GBD) 2021 study further highlight its substantial global impact, reporting 107.6 million prevalent cases of Diabetic Nephropathy worldwide and 477,300 deaths attributable to the disease in 2021 (Zhang et al., 2025). DN primarily arises from chronic hyperglycemia, which initiates a cascade of pathological events including oxidative stress, persistent inflammation, and the accumulation of advanced glycation end products ultimately resulting in progressive renal tissue damage and functional decline (Mima, et al., 2025). Collectively, these processes reflect sustained metabolic dysregulation, wherein prolonged metabolic imbalance disrupts cellular homeostasis and accelerates organ-specific injury (Yamagishi & Matsui, 2010; Martínez-López et al., 2025).

At the molecular level, accumulating evidence from transcriptomic studies has established that aberrant gene expression is a central driver of metabolic disorder progression (Liu et al., 2025). In particular, upregulated genes are frequently enriched in key biological pathways such as immune and inflammatory signaling, extracellular matrix remodelling, and cellular stress responses (Tang & Yiu, 2020). In the context of DN, these transcriptional alterations not only serve as molecular indicators of disease activity but also actively contribute to renal dysfunction and structural damage, underscoring their relevance as potential therapeutic targets (Tang & Yiu, 2020; Samsu, 2021). Because the downstream objective of this study is to perform virtual screening for putative receptor antagonists (inhibitors), our analysis deliberately focuses on these upregulated genes. These overexpressed genes represent hyperactive pathological cascades that can be therapeutically suppressed by inhibitory small molecules, whereas downregulated transcripts typically represent lost protective mechanisms requiring gene-restoration or agonistic therapies outside the current scope. Advances in RNA sequencing (RNA-Seq) technologies have further enabled high-resolution, genome-wide gene expression profiling in disease conditions (Martínez-López et. al, 2025; Singh et al., 2026a). Genome-wide differential gene expression studies on DN reveal complex networks of dysregulated genes that orchestrate disease progression and provide mechanistic insights into renal injury (Ma et al., 2025). Consequently, a

substantial volume of DN-related transcriptomic datasets and analytical resources are now publicly available, offering unprecedented opportunities for in-depth computational analyses (Cai et al., 2025).

To comprehensively interpret large-scale transcriptomic datasets, integrative computational and systems-level approaches have gained increasing prominence. In particular, integrated methods centred on molecular network analysis combined with machine learning and other advanced data-mining strategies have emerged as powerful tools for extracting biologically meaningful insights from complex omics data (Verma et al., 2020; Wen et al., 2025). When integrated with computational drug discovery, such approaches enhance the prioritization of candidate therapeutic targets and facilitate the development of more precise and personalized interventions for complex diseases (Zhou et al., 2018; Tang & Yiu, 2020). Recent studies further underscore the translational potential of these integrated methods across diverse pathological conditions, including Periodontitis, diabetes in murine models, and viral respiratory tract infections (Wang et al., 2024; Wen et al., 2025; Tao et al., 2025). Accordingly, the application of these *in silico* approaches may also provide valuable insights into the pathogenesis of DN and support the identification of novel therapeutic opportunities.

Owing to the adverse effects and long-term safety concerns associated with many synthetic compounds, the therapeutic use of natural compounds has gained increasing attention in recent years (Rana et al., 2022). The therapeutic relevance of these natural agents is supported by both experimental and clinical studies demonstrating their ability to simultaneously modulate multiple pathological processes, including oxidative stress, inflammatory cytokine production, and renal fibrosis; they are also capable of regulating gene expression and intracellular signaling cascades. Metabolic disorders, including DN, are no exception to this paradigm. A growing body of evidence indicates that plant-derived bioactive compounds such as flavonoids, polyphenols, alkaloids, terpenoids, glycosides, and polysaccharides can attenuate DN progression by modulating key molecular pathways associated with oxidative stress, inflammation, and extracellular matrix (Aswal et al., 2020; Hu et al., 2023; Cotas et al., 2024; Li et al., 2025). Although the therapeutic potential of these compounds has been substantiated through experimental studies, their clinical translation as target-specific therapeutics remains limited, largely due to insufficient molecular-level understanding of their mechanisms of action, target specificity, and binding affinities.

In summary, the interplay between metabolic dysregulation, aberrant gene expression, oxidative stress, and inflammation forms the molecular foundation of DN progression. We hypothesize

that hyperactive disease modules in DN can be identified through transcriptomic network analysis and therapeutically targeted using multi-target natural antagonists. To explore this hypothesis and generate robust, data-driven hypotheses for future experimental validation, this study pursues three distinct objectives. First, we aim to identify robust disease-associated gene modules by analyzing upregulated differentially expressed genes from human DN RNA-Seq data. Second, we seek to construct and elucidate candidate 'Gene–miRNA–TF' regulatory networks to understand how microRNAs and transcription factors stabilize these hyperactive disease programs. Finally, we computationally prioritize key molecular targets from these networks to assess their theoretical druggability against natural compounds via molecular docking, ultimately prioritizing candidate multi-target antagonists.

2. MATERIALS & METHODOLOGY

2.1 Data Retrieval

RNA-seq–based gene expression datasets related to DN were retrieved from the European Nucleotide Archive (ENA) under BioProject PRJNA595590 (also accessible via GEO: GSE142025) (Table 1). The dataset comprises human kidney biopsy samples with clearly defined control and diabetic nephropathy groups, generated using paired-end Illumina sequencing technology (Fan et al., 2019, 2020).

Table 1: Description of the RNA-Seq dataset used in this study

Group	Samples	Run Accessions	Sample IDs (SAMN)	GEO Accessions
Control	**9**	SRR10691658-66	SAMN13560918-29	GSM4217808-16
Advanced DN	**21**	SRR10691631-51	SAMN13560910-17,34-46	GSM4217781-7801

2.2 Data Preprocessing

Raw FASTQ files were initially assessed using FastQC to evaluate per-base quality scores, GC content, sequence duplication levels, and adapter contamination (Andrews, 2010). Adapter sequences and low-quality bases were subsequently trimmed using Trimmomatic for paired-end reads. Trimming was performed using the following strict parameters to ensure high-quality inputs: LEADING:20 TRAILING:20 SLIDINGWINDOW: 4:20 MINLEN:20. To clarify the quality threshold and correct the Phred+33 encoding nomenclature, reads were filtered based on a true Q-score cutoff of $Q \geq 20$. Only reads meeting these predefined thresholds were retained for downstream alignment.

2.3 Differential Gene Expression Analysis

To ensure robust, splice-aware mapping of the transcriptomic data, raw sequencing paired-end reads were aligned to the human reference genome (GRCh38) using HISAT2. Following alignment, sorting, and indexing of the BAM files via SAMtools, downstream gene quantification was performed using the *featureCounts* function within the Rsubread package. Reads were quantified against the GRCh38 human reference genome annotation, using a protein-coding GTF annotation file utilizing the following explicit parameters: GTF.featureType = "exon", GTF.attrType = "gene_id", isPairedEnd = TRUE, and strandSpecific = 0 (unstranded). Raw counts were normalized for library size and compositional bias using DESeq2 (Love et al., 2014). Differential expression analysis was conducted, and genes with an adjusted p-value ≤ 0.05 and an absolute $\log_2$ fold change ≥ 1 were considered DEGs. Because the downstream objective relies on identifying small-molecule antagonists to suppress hyperactive pathological cascades, only significantly upregulated genes ($\log_2$ fold change > 1, adjusted $p \leq 0.05$) were retained for subsequent clustering and network analyses. Downregulated transcripts, which typically reflect lost protective mechanisms requiring agonistic or gene-restoration therapies, were deliberately excluded as they fall outside the scope of the virtual inhibitory screen.

2.4 Identification of Disease-Associated Gene (DAG) Modules

To define robust co-expression modules, upregulated DEGs were clustered using K-means clustering (seed = 123; nstart = 25). To assess module topology, graph clustering coefficients were calculated using the igraph package, retaining only robust clusters (average coefficient > 0.7). These clusters were then intersected with 247 established metabolic drug targets (DrugBank) to prioritize disease-relevant modules. Co-expression within these modules was validated using Pearson correlation ($|r| \geq 0.7$) with Benjamini-Hochberg FDR correction (FDR ≤ 0.05). These high-confidence genes were defined as disease-associated genes (DAGs) and served as the input for network reconstruction.

2.5 Network Reconstruction

2.5.1 Protein–protein interaction (PPI) network

The DAGs were mapped to the STRING database v12.0 to retrieve high-confidence protein–protein interactions (score ≥ 0.70) to ensure network robustness and minimize false-positive edges. The resulting network was visualized in Cytoscape. Following network visualization, the *cytoHubba* plugin for Cytoscape (Chin et al., 2014) was employed to rank nodes based on degree and betweenness centrality, enabling the objective identification of central hub proteins.

Nodes within the top 10–20% of either metric were classified as hub or bottleneck proteins, highlighting key regulators of network connectivity.

2.5.2 Prioritization of Candidate Docking Targets

To bridge the systems-biology findings to therapeutic discovery, 29 specific protein targets were prioritized from the reconstructed PPI network. These 29 targets were selected based on three criteria: (1) high degree centrality (hub status) within the PPI network, (2) identification as core enriched genes in the inflammatory and fibrotic GO pathways, and (3) the availability of high-resolution (≤ 2.5 Å) crystal structures in the Protein Data Bank (PDB). A mapping of these targets to their PDB structures and the rationale for their selection is provided in (Table S1).

2.5.3 Regulatory Network Analysis

To elucidate gene regulatory relationships, hub genes were submitted to miRNet 2.0 (Fan et al., 2016; Chang et al., 2020). To avoid non-specific network density, interactions were strictly filtered using the following parameters: species = *Homo sapiens*. The resulting regulatory network, integrating TF–gene and miRNA–gene interactions, was visualized to identify regulators targeting multiple hub genes, which were prioritized as candidate master regulators and potential intervention points.

2.6 Gene Enrichment Analysis

Functional enrichment analysis was performed in R using the *clusterProfiler* 4.0 package (Wu et al., 2021) via over-representation analysis (ORA). Gene Ontology (GO) and KEGG pathway enrichments were computed using Benjamini–Hochberg adjusted *p*-values ≤ 0.05. To maintain biological context, enrichment terms were analyzed with the understanding that in bulk biopsy datasets, immune-related pathways (e.g., leukocyte activation) reflect both intrinsic renal responses and the presence of infiltrating immune cells characteristic of DN.

2.7 Natural Compound Selection and Preparation

Plant-derived natural products from *Tracheophyta* species with reported antidiabetic relevance were retrieved from the LOTUS database (Gaonkar & Hullatti, 2020). The filtering workflow was executed in nine consecutive stages to refine the initial dataset for high-confidence, bioactive molecules. Beginning with an initial retrieval of 276,518 compounds, the dataset was first processed by removing records containing missing (NA) values, which retained 236,364 entries. This was followed by a biological relevance filter that restricted the dataset to compounds from 36 families and 50 species specifically relevant to antidiabetic activity,

resulting in 2,744 retained rows. After removing rows with empty strings or missing data to ensure data integrity, 2,518 candidates remained. We then applied a Natural Product-Likeness (NPL) score filter of ≥ 1 to prioritize structural scaffolds typical of secondary metabolites, retaining 2,207 candidates. Duplicate compounds were identified and removed based on InChIKey, maintaining 2,207 unique entries. We subsequently applied a molecular weight filter to retain drug-like molecules between 150–600 Da, resulting in 1,438 candidates, followed by the application of Lipinski's Rule of Five (MW ≤ 500 Da, XLogP ≤ 5, H-bond donors ≤ 5, H-bond acceptors ≤ 10), which yielded 830 candidates. In the final refinement stage, these 830 candidates were filtered to exclude compounds not representative of established bioactive classes specifically removing fatty alcohols, wax monoesters/diesters, triacylglycerols, long-chain hydrocarbons, simple amino acids, and primary metabolites while prioritizing bioactive-rich classes such as flavonoids, limonoids, alkaloids, terpenoids, lignans, phenolics, and coumarins. This rigorous process yielded a final high-confidence library of 640 candidate compounds (Table S2).

Structural preparation of these 640 candidates was performed using RDKit to optimize molecular geometries. Compounds were converted to PDBQT format using Meeko, utilizing the following parameters: --addhydrogens, --partialcharge gasteiger, and --torsions. Ligand protonation and tautomer states were standardized at physiological pH 7.4. Target protein structures were retrieved from the Protein Data Bank (PDB; Berman et al., 2000), curated in PyMOL to remove non-essential heteroatoms (e.g., water molecules and co-crystallized buffers), and prepared using AutoDock Tools/MGLTools version 1.5.6. Throughout the virtual screening process, receptors were treated as rigid bodies to maintain the structural integrity of the active sites.

2.8 Molecular Docking and Screening

Blind molecular docking was performed using AutoDock Vina version 1.2.5 (Eberhardt et al., 2021). To ensure a comprehensive, unbiased, and reproducible search of potential binding sites, the docking grid parameters were dynamically calculated for each individual receptor. The grid box center for each target was defined by the geometric center of the protein's atomic coordinates, and the grid box dimensions were automatically generated to encompass the entire protein surface with an additional 20 Å of padding, ensuring full target coverage (Table S3). This automated, protein-specific grid generation strategy eliminates subjective bias in binding site identification and ensures consistent search space coverage across the diverse structural library. We utilized an exhaustiveness of 8 and a CPU count of 8 to ensure sufficient

conformational sampling. Compounds with a binding affinity ≤ −8 kcal/mol were prioritized for post-docking analysis. Results were visualized using PyMOL and Discovery Studio.

2.9 Post-Docking Network Analysis

To evaluate the polypharmacological potential of the prioritized natural compounds, a protein–ligand interaction network was constructed using Cytoscape. In this bipartite network, protein targets and ligands were represented as nodes, while docking-derived associations were depicted as edges, with weights assigned based on the predicted binding affinity (kcal/mol). The network topology was interrogated to identify high-degree ligands compounds interacting with multiple protein targets and key protein hubs. Ligands engaging with two or more disease-associated targets were classified as multi-target candidates, providing a system-level perspective on the potential for pathway-level modulation in diabetic nephropathy.

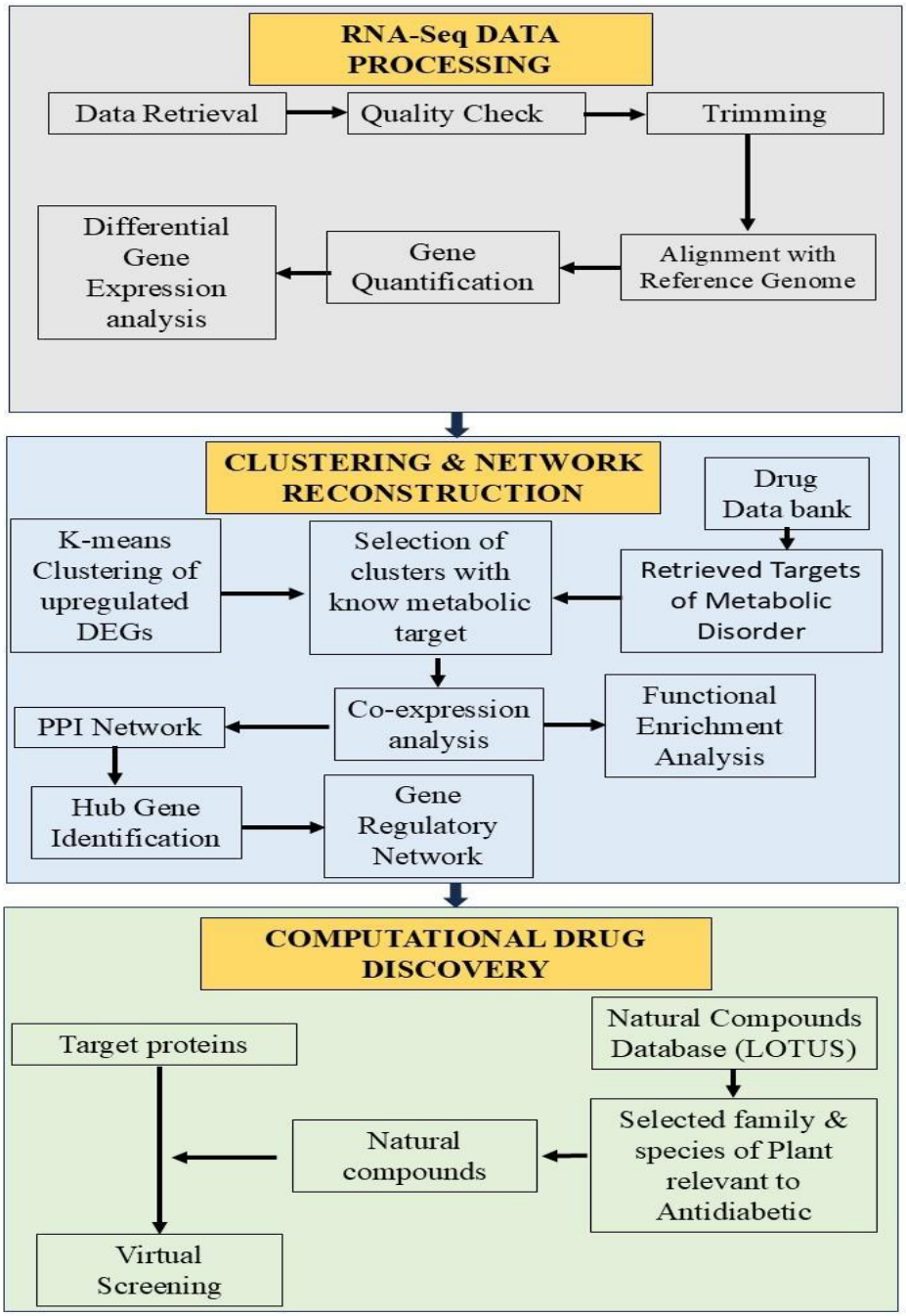


Figure 1: Overview of the integrated computational workflow illustrating data processing, network reconstruction and analysis, and molecular docking

3. RESULTS:

3.1 Differential Gene Expression

Differential expression analysis performed using DESeq2 identified 2,075 DEGs between DN and control samples (adjusted $p \leq 0.05$; $|\log_2 \text{fold change}| \geq 1$), of which 1,484 were upregulated and 591 were downregulated in DN (Love et al., 2014). The complete list of significant DEGs, including $\log_2$ fold change values and adjusted p values, is provided in Table S4.

The volcano plot depicts the global distribution of DEGs, highlighting a predominance of upregulated genes with larger positive $\log_2$ fold changes, which therefore constituted the primary focus of subsequent downstream analyses (Figure 2).

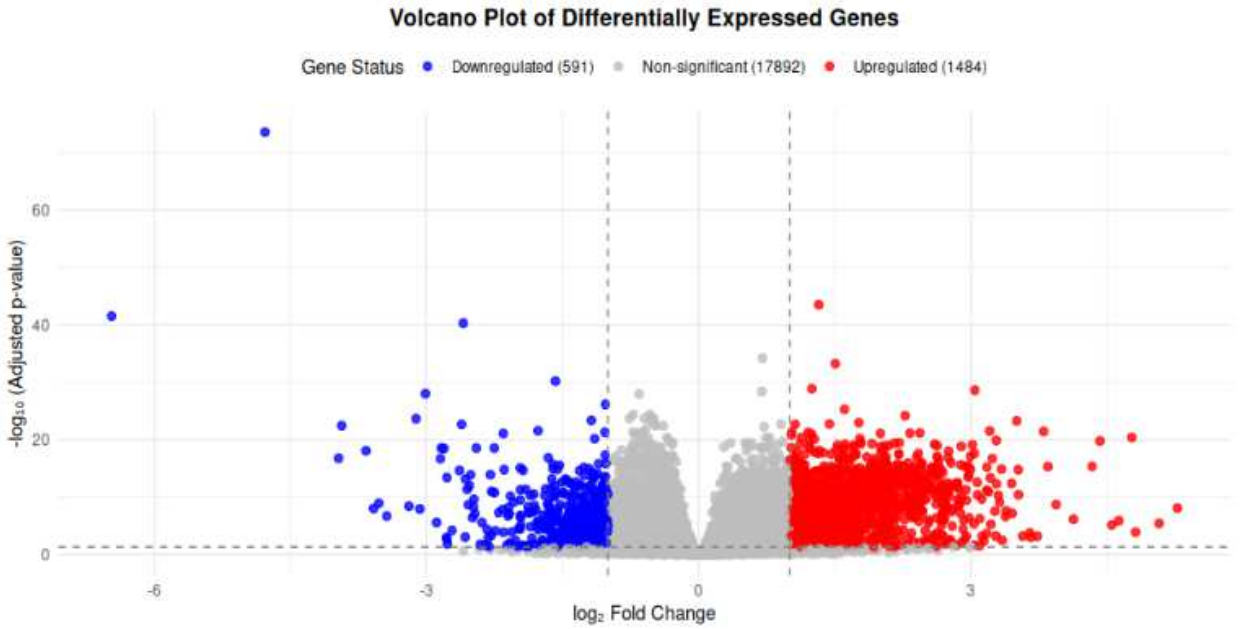


Figure 2: Volcano plot showing the distribution of differentially expressed genes according to $\log_2$ fold change and $-\log_{10}$ adjusted p-value

3.2 Cluster Properties and Drug Target Overlap

Stringent K-means clustering (k =75) was subsequently applied to the upregulated DEGs, resulting in 75 distinct expression clusters (Table S5). For each cluster, both the cluster size (number of genes) and the number of annotated DrugBank targets were quantified. Notably, larger clusters tended to harbor a greater number of metabolic drug targets, suggesting enhanced functional and therapeutic relevance in the context of diabetic nephropathy (Wishart et al., 2018).

To assess network robustness, clustering coefficients were calculated for all clusters. Fourteen clusters exceeded the predefined robustness threshold (clustering coefficient ≥ 0.7), indicating strong internal connectivity and the presence of tightly coordinated gene modules. Clusters that simultaneously contained at least one known metabolic drug target and exhibited high clustering coefficients were therefore designated as biologically relevant modules. Detailed characteristics of these modules including cluster size, mean clustering coefficient and DrugBank target counts are summarized in Table S6.

The number of strong co-expression pairs ($|r| \geq 0.7$) within each cluster was quantified, and clusters harbouring the highest numbers of such interactions were prioritized for downstream network analyses, as they represent tightly coordinated gene modules in diabetic nephropathy (Ruan et al., 2010). These findings are summarized in Figure 3, while the complete pairwise correlation counts and cluster-level summary statistics are provided in Table S7.

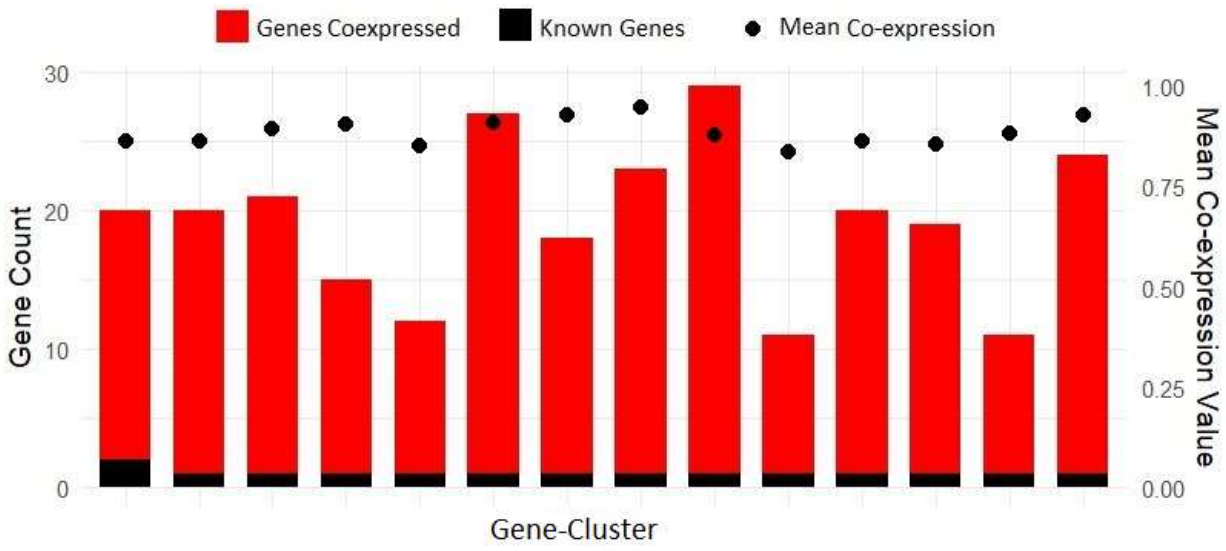


Figure 3: Graph summarizing gene counts and co-expression pairs within the selected clusters of interest used for network reconstruction and downstream analyses.

3.3 Network reconstruction

3.3.1 Protein–protein interaction network

Three hundred and thirty-six co-expressed genes retained from prioritized clusters were mapped to the STRING database to construct a protein–protein interaction (PPI) network. PPI network analysis was restricted to experimentally supported, high-confidence interactions (Szklarczyk et al., 2025). The initial STRING network comprised 329 nodes and 277 edges. This network was subsequently imported into Cytoscape for further refinement and analysis. To enhance biological relevance and network robustness, proteins involved in only a single interaction were removed, resulting in a filtered PPI network containing 104 nodes and 230 edges. The resulting network was visualized and interrogated using Cytoscape, which revealed densely interconnected modules enriched for upregulated genes associated with diabetic nephropathy.

This PPI network delineates key functional interactions among disease-relevant genes and provides a robust framework for the identification of central regulatory nodes and putative therapeutic targets. The overall network architecture and major modules are presented in Figure

4, while the corresponding STRING interaction file and Cytoscape node attribute table are provided as Supplementary tables (Table S8 & S9).

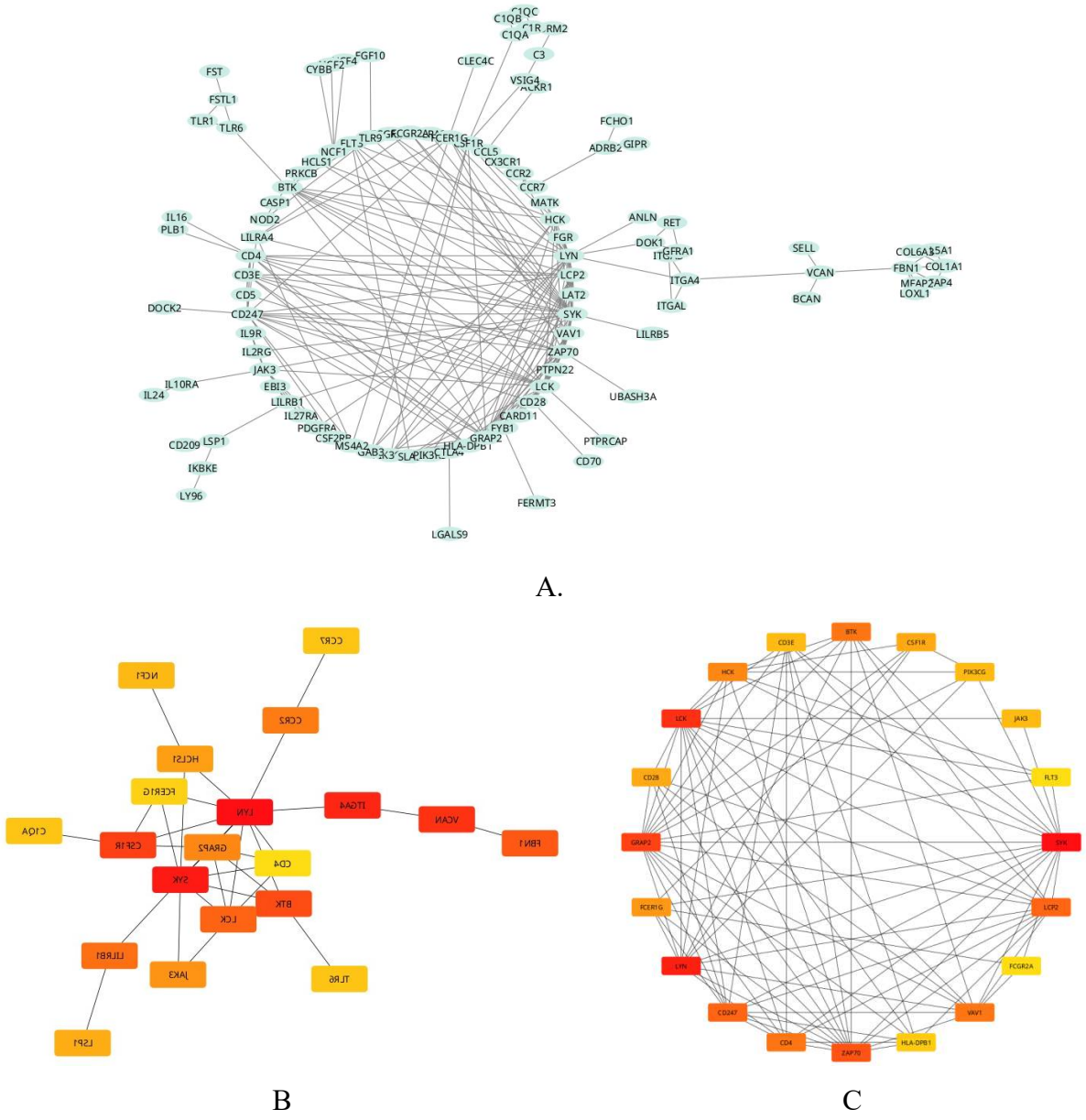


Figure 4: Protein–protein interaction (PPI) network of co-expressed upregulated genes in diabetic nephropathy. (A) Global PPI network constructed from significantly co-expressed upregulated genes, illustrating the overall interaction landscape. (B) Subnetwork highlighting the top 20 hub genes ranked by betweenness centrality, reflecting key nodes involved in information flow across the network. (C) Subnetwork highlighting the top 20 hub genes ranked by degree centrality**,** representing highly connected proteins within the PPI network

Network topology was then quantified using degree and betweenness centrality, and the cytoHubba plugin was applied to rank all nodes according to these metrics (Chin et al., 2014). The top 20 nodes by degree and the top 20 nodes by betweenness centrality were designated as

key proteins for target prioritization (Figure 4b, c), and their gene symbols and centrality scores are listed in Table 2.

Table 2: List of genes of importance with their degree and betweenness

S.N.	Centrality	Gene list with centrality values
1	Degree	*SYK* (25), *LYN* (24), *LCK* (20), *GRAP2* (19), *ZAP70* (17), *LCP2* (15), *CD247* (15), *BTK* (12), *VAV1*(12), *CD4* (12), *HCK* (11), *FCER1G* (10), *CSF1R* (9), *CD28* (9), *CD3E* (8), *PIK3CG* (8), *JAK3* (8), *HLA-DPB1* (7), *FCGR2A* (6), *FLT3* (6)
2	Betweenness	*LYN* (0.467681), *SYK* (0.264722), *ITGA4* (0.211689), *VCAN* (0.163906), *CSF1R* (0.162569), *BTK* (0.131715), *FBN*1 (0.112253), *LCK* (0.095055), *LILRB1* (0.09205), *CCR2* (0.08569), *GRAP2* (0.082776), *JAK3* (0.077205), *HCLS1* (0.064563), *LSP1*(0.057491), *NCF1* (0.057301), *CCR7* (0.05711), *TLR6* (0.05711), *C1QA* (0.05711), *FCER1G* (0.054901), *CD4* (0.049018)

3.4 Regulatory Network Analysis

Following reconstruction of the PPI network, a gene regulatory network was generated for the identified hub genes to delineate potential miRNAs and TFs regulators associated with diabetic nephropathy. Hub genes were submitted to miRNet database to retrieve experimentally validated and predicted 'miRNA–gene' and 'TF–gene' interactions (Fan et al., 2016; Chang et al., 2020). The resulting regulatory network reveals extensive connectivity between miRNAs, TFs, and multiple hub genes, indicating their potential roles as central regulatory elements in disease-associated gene regulation (Figure 5).

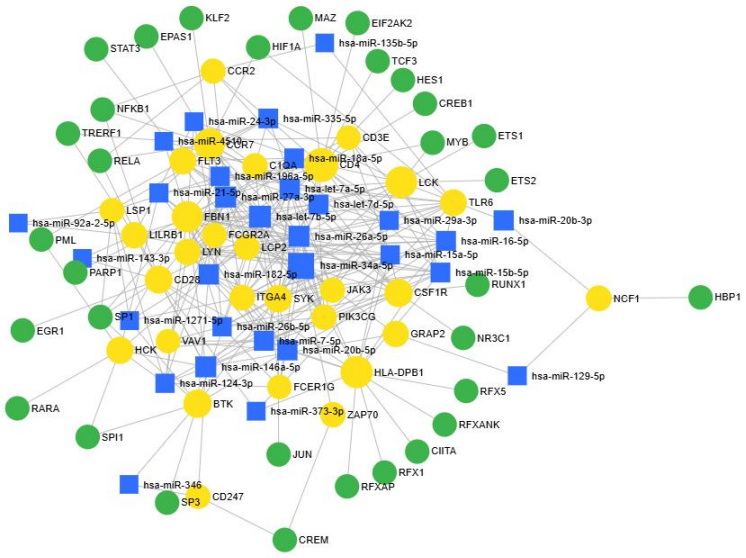


Figure 5: This regulatory network highlights key TFs (green), miRNAs (blue) and Genes (yellow) as master regulators in diabetic nephropathy.

Our analysis integrated 'TF–gene' and 'miRNA–gene' interactions to construct a comprehensive regulatory network, highlighting the connectivity of key regulatory elements in the context of diabetic nephropathy.

3.4.1 Transcription Factor-Gene Interactions

The 'TF–gene' regulatory network identified 32 TFs mapped to their corresponding target genes in diabetic nephropathy, including prominent regulators such as *NFKB1, RELA, SP1, and STAT3*. Collectively, these TFs were linked to the regulation of multiple disease-relevant genes, notably *CCR7, CD28, LCK, and CSF1R*, indicating that a relatively limited set of TFs governs a broad immune and signaling transcriptional program in the diabetic kidney.

NFKB1 and RELA, the canonical subunits of the NF-κB signaling pathway, were associated with chemokine receptors and costimulatory molecules, consistent with their well-established roles in mediating inflammatory and immune-responsive gene expression in renal and immune cell populations (Mezzano et al., 2004; Navarro-González et al., 2011). *SP1*, a GC-box–binding zinc finger transcription factor, exhibited extensive connectivity across both immune-related and structural genes, suggesting a potential role in integrating hyperglycemia-induced cellular stress with core transcriptional regulation (Reddy and Natarajan, 2013). *STAT3*, a cytokine-activated signal transducer and transcriptional regulator, emerged as another major hub, linking cytokine

signaling pathways to downstream genes involved in cell activation, survival, and myeloid receptor signaling (Wang et al., 2012).

Taken together, these 32 TFs constitute a densely interconnected regulatory layer in diabetic nephropathy, with *NFKB1, RELA, SP1,* and *STAT3* functioning as principal hubs coordinating the expression of chemokine receptors, T-cell signaling components, and myeloid lineage markers. The complete set of curated TF–gene interactions supporting this regulatory network is provided in Table S10.

3.5 miRNA-Gene Interactions

The 'miRNA–gene' regulatory network identified a total of 683 miRNAs, including well-characterized regulators such as *hsa-let-7a-5p, hsa-miR-15a-5p, hsa-miR-16-5p, hsa-miR-21-5p,* and *hsa-miR-29a-3p*, targeting genes such as *CD4, CCR7, CSF1R, FBN1, ITGA4, LCK, SYK,* and *TLR6* consistent with macrophage-driven inflammation and fibrosis in diabetic nephropathy (Dong et al.,2024). Many of these miRNAs are established modulators of inflammation, immune signaling, and extracellular matrix (ECM) remodelling, supporting their functional relevance in metabolic and inflammatory disorders, including diabetic nephropathy. The complete set of miRNA–gene interactions comprising this network is provided in Table S11.

Several hub genes, including *CD4, CCR7, CSF1R, FBN1, ITGA4, LCK, SYK*, and *TLR6*, were regulated by multiple transcription factors and miRNAs, underscoring their central regulatory positioning within the network. These genes play key roles in immune cell activation, leukocyte signaling, and ECM organization, further implicating them in the molecular pathology of diabetic complications.

Network motif analysis revealed a significant enrichment of feed-forward loop (FFL) structures, in which a TF and a miRNA jointly regulate shared target genes. Such motifs represent robust regulatory modules capable of fine-tuning gene expression and may contribute to disease-specific transcriptional programs observed in diabetic nephropathy.

Overall, this integrated 'gene-miRNA–TF' regulatory network provides a systems-level framework for understanding the complex regulatory architecture underlying diabetic nephropathy, highlighting key regulatory hubs and pathways that may represent potential targets for therapeutic intervention.

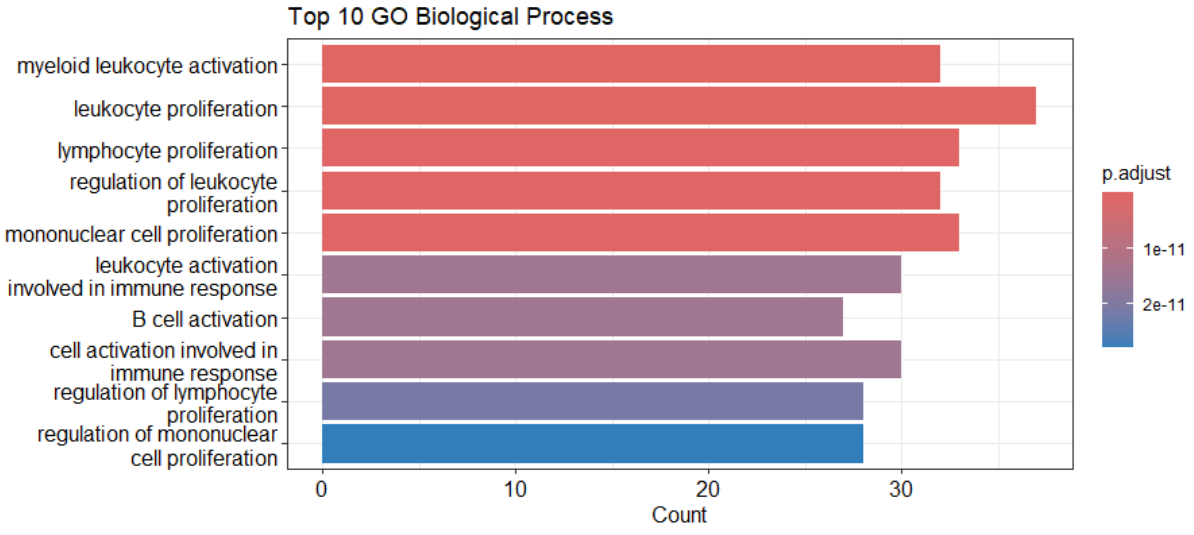
3.6 Gene Enrichment Analysis

To delineate the functional landscape associated with the disease-related gene set, functional enrichment analysis was performed using all genes retained from the filtered co-expression network. Official HGNC-approved human gene symbols corresponding to these network genes were used as input for over-representation analysis (ORA) of Gene Ontology (GO) categories—biological process (BP), molecular function (MF), and cellular component (CC)—as well as Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways (Ashburner et al., 2000).

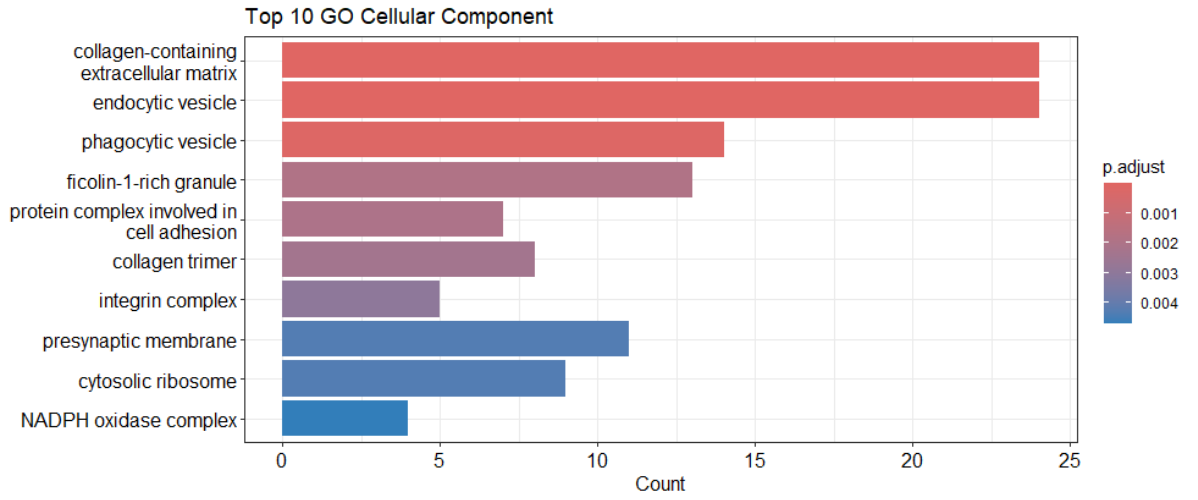
Enrichment testing was conducted using a hypergeometric statistical framework, with Benjamini–Hochberg multiple testing correction applied to control the false discovery rate (FDR). Terms with an adjusted p-value ≤ 0.05 were considered significantly enriched. This analytical approach enabled the systematic identification of biological processes, molecular activities, cellular localizations, and signaling pathways that are overrepresented among disease-associated genes.

The enrichment results were summarized and visualized as bar plots (Figure 6), displaying the top 10 most significantly enriched terms for each GO category (BP, CC, MF, and KEGG pathways). Within each panel, terms were ranked by adjusted p-value, thereby emphasizing the most statistically robust and biologically relevant functional themes linked to the disease-associated gene set. Collectively, these enrichment profiles highlight key immune, signaling, and cellular structural pathways that may contribute to disease pathophysiology.

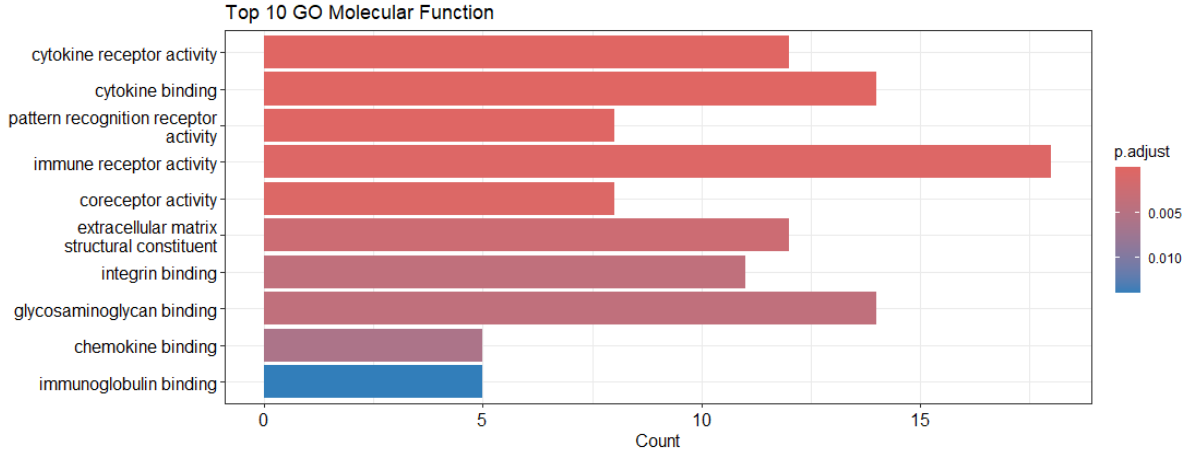
Complete enrichment output tables, including enriched term descriptions, gene counts, associated gene lists, enrichment statistics, and adjusted p-values, are provided in Tables S12–S15, enabling full transparency and reproducibility of the functional analysis.



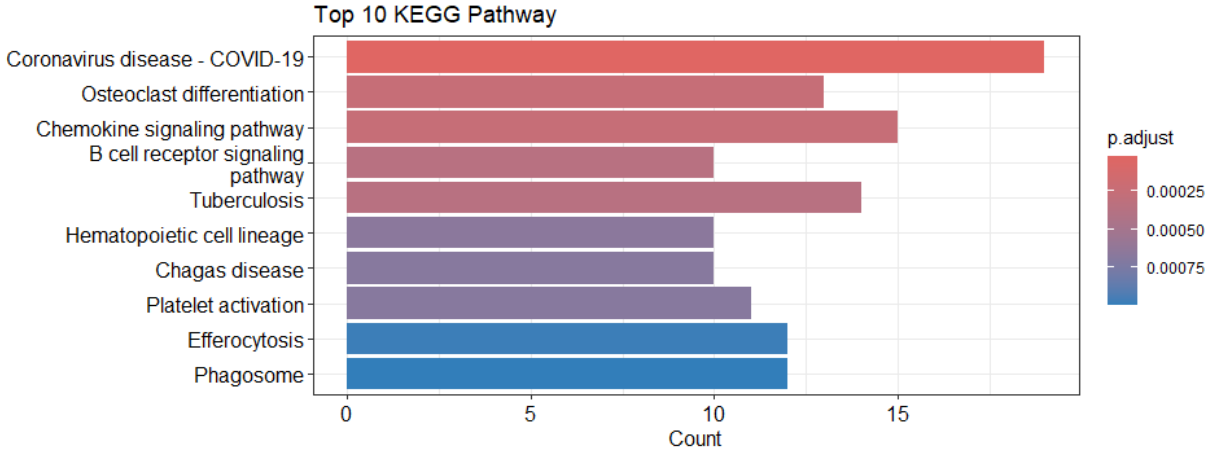
(A)



(B)



(C)



(D)

Figure 6: Functional enrichment of disease-associated network genes. (A) Top 10 significantly enriched Gene Ontology biological process terms. (B) Enriched GO cellular component terms. (C) Enriched GO molecular function terms. (D) Significantly enriched Kyoto Encyclopedia of

Genes and Genomes (KEGG) pathways. Bars are ranked by adjusted p-value, with the most significant terms shown first.

Biological Process (BP) enrichment analysis revealed strong over-representation of immune cell activation and proliferation processes, consistent with inflammatory mechanisms in diabetic nephropathy (Figure 6). Prominently enriched terms spanned leukocyte activation (e.g., myeloid leukocyte activation, leukocyte activation involved in immune response, B cell activation, cell activation involved in immune response) and proliferation (e.g., leukocyte proliferation, lymphocyte proliferation, mononuclear cell proliferation, regulation of leukocyte proliferation, regulation of lymphocyte proliferation, regulation of mononuclear cell proliferation).

Cellular Component (CC) enrichment similarly highlighted extracellular matrix and vesicular terms, underscoring fibrotic remodelling and phagocytic activity (e.g., collagen-containing extracellular matrix, endocytic vesicle, phagocytic vesicle, ficolin-1-rich granule, protein complex involved in cell adhesion, collagen trimer, integrin complex, presynaptic membrane, cytosolic ribosome, NADPH oxidase complex).

In Molecular Function (MF) ontology, cytokine/chemokine and immune receptor activities are predominated, reflecting role of genes in intensified inflammatory signaling (e.g., cytokine receptor activity, cytokine binding, pattern recognition receptor activity, immune receptor activity, coreceptor activity, extracellular matrix structural constituent, integrin binding, glycosaminoglycan binding, chemokine binding, immunoglobulin binding).

KEGG pathway analysis corroborated these findings with over-representation of immune signaling and phagocytic pathways (e.g., Coronavirus disease [COVID-19], Osteoclast differentiation, Chemokine signaling pathway, B cell receptor signaling pathway, Tuberculosis, Hematopoietic cell lineage, Chagas disease, Platelet activation, Efferocytosis, Phagosome), indicative of dysregulated inflammation and pathogen-like responses.

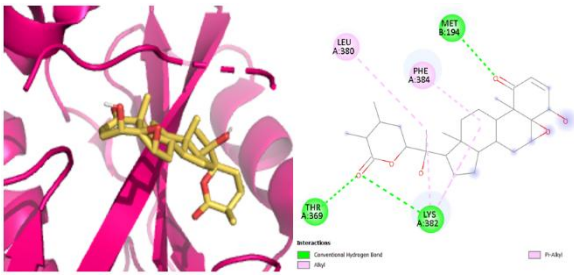
3.7 Molecular Docking Analysis and Selection of Protein–Ligand Hits

3.7.1 Screening of Natural Compounds

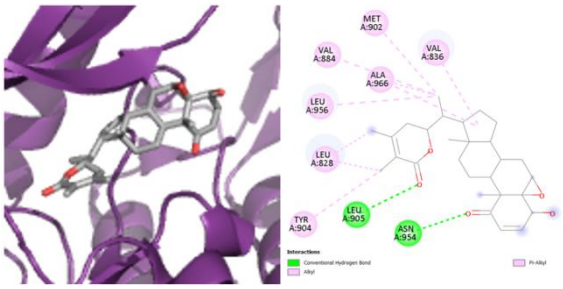
Molecular docking analysis was conducted using AutoDock Vina (Trott & Olson, 2010) to evaluate the binding potential of retrieved natural compounds against multiple protein targets implicated in diabetic nephropathy. In total, 18,502 docking simulations were successfully completed and analyzed (the remaining 58 simulations were excluded due to structural preparation errors or convergence failures during the docking process, representing a <0.4%

exclusion rate). For each protein–ligand pair, the top-ranked binding pose (Mode 1) corresponding to the lowest predicted binding free energy was extracted. The associated binding affinity values (kcal/mol) were used as the primary quantitative criterion for ranking and prioritizing candidate ligands.

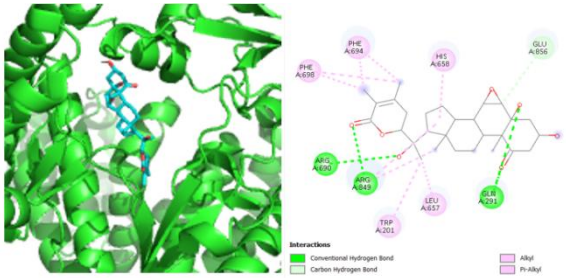
This high-throughput docking strategy enabled a comprehensive assessment of ligand–target interaction landscapes across the selected protein targets, providing an initial in silico filter to identify high-affinity natural compounds with potential relevance to diabetic nephropathy–associated pathways. The resulting binding energy profiles served as the basis for subsequent comparative analysis and downstream validation steps, including interaction profiling and candidate prioritization.



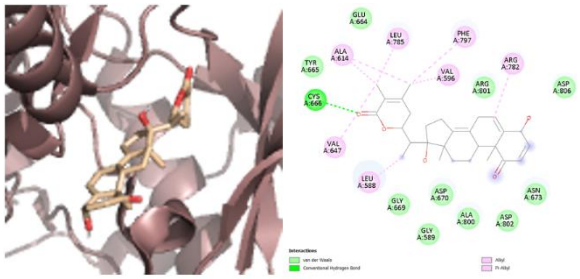
5GJI-LTS0080198



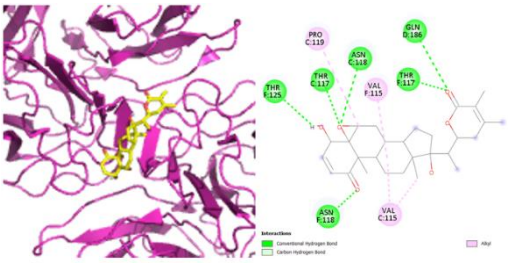
5LWM-LTS0251312



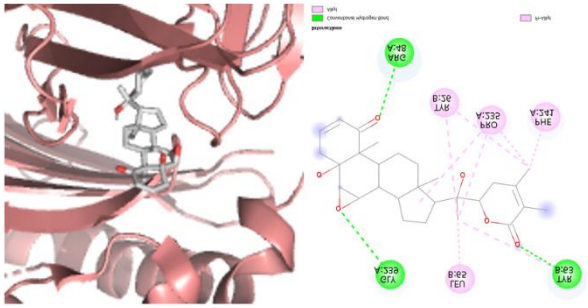
1E8Y-LTS0168926



8JOT-LTS0206640



2WNV-LTS0095642



3BH8-LTS0153288

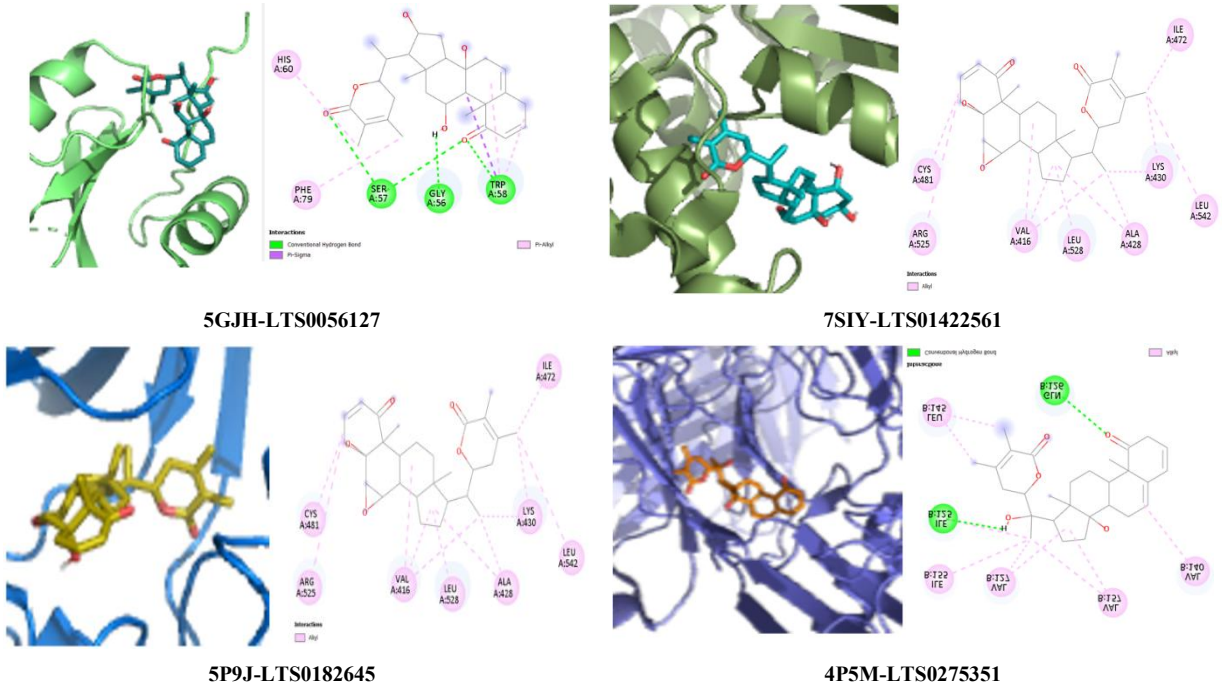


Figure 7: Two- and three-dimensional visualizations of the highest-affinity docked complexes (top 10) generated using molecular visualization tools, highlighting atom-level ligand–protein interactions.

3.7.2 Selection of Top-Binding Ligands for Each Protein Target

For each protein target, docked ligands were ranked according to their predicted binding affinities, with more negative binding energy values indicative of stronger ligand–target interactions. To enable standardized and unbiased comparison across all targets, the top 10 highest-affinity ligands were selected for each protein. This uniform ranking strategy minimized target-specific selection bias while retaining the most promising candidates per receptor.

Application of this filtering criterion yielded a curated high-confidence dataset consisting of 10 ligands per target across 29 protein targets, resulting in a total of 290 prioritized protein–ligand interactions. This refined interaction set provided a robust foundation for downstream comparative analyses, including binding mode inspection, interaction profiling, and subsequent in silico validation.

3.7.3 Construction of the Protein–Ligand Interaction Network

To explore the potential for polypharmacological interactions between predicted ligands and their protein targets, a network pharmacology framework was employed. To systematically characterize the interaction landscape of the prioritized docking hits, a protein–ligand interaction network was constructed and visualized using Cytoscape. In this bipartite network, protein

targets and ligands were represented as nodes, while docking-derived protein–ligand associations were depicted as edges weighted by predicted binding affinity, with binding energy values (kcal/mol) assigned as edge weights.

Network visualization and interrogation enabled the identification of highly connected ligands exhibiting interactions with multiple protein targets, as well as key proteins that engaged with several high-affinity compounds, thereby highlighting potential multi-target ligands and interaction hubs within the docking landscape. Such connectivity patterns provide insight into the polypharmacological potential of candidate compounds and the relative target promiscuity or selectivity of individual proteins.

The resulting node and edge attribute tables constitute a comprehensive dataset that supports network-based visualization, topological analysis, and comparative prioritization of protein–ligand interactions, and can be readily integrated with additional molecular or systems-level analyses.

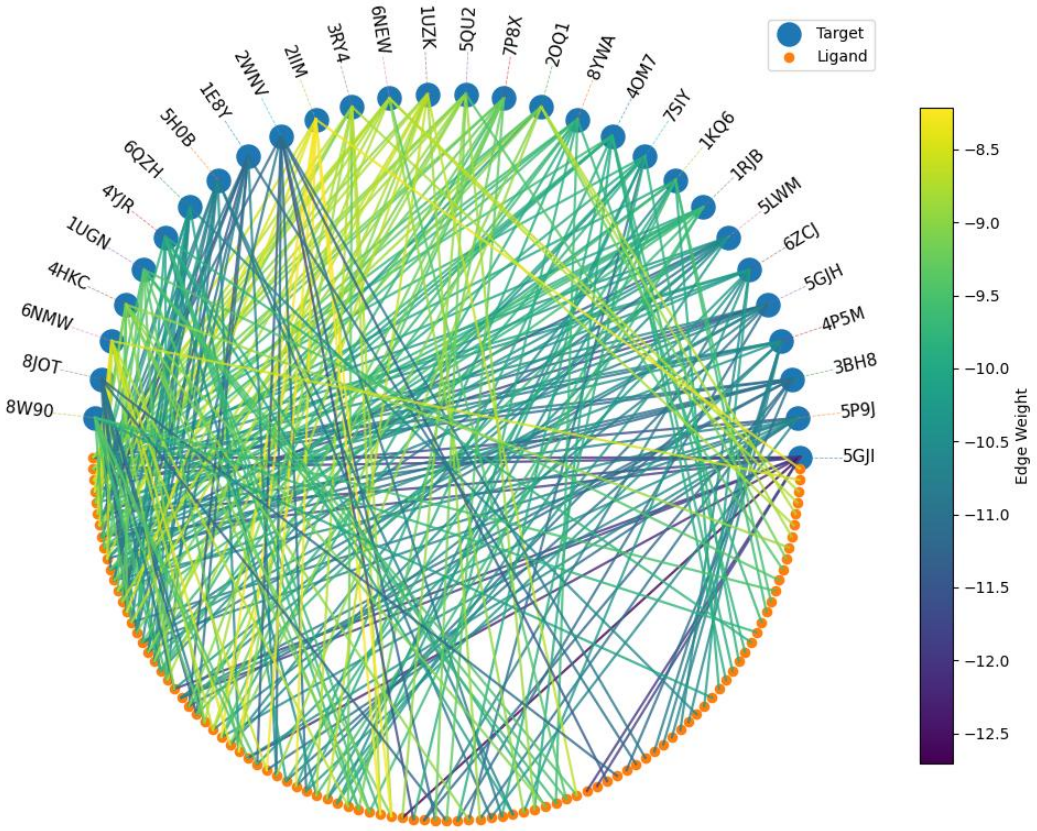


Figure 8: Protein–Ligand Interaction Network of Top 10 docked ligand for each protein. A circular bipartite network illustrating interactions between protein targets and small-molecule ligands. Protein targets are arranged along the upper semicircle and shown as larger nodes, whereas ligands are positioned along the lower semicircle and represented as small dots. Edges

denote ligand–target interactions, with edge colors reflecting interaction strength according to the continuous color scale (low to high weight).

3.7.4 Identification of Multi-Target Ligands with Polypharmacological Potential

Based on the protein–ligand interaction network, ligands exhibiting interactions with more than one protein target were classified as multi-target ligands (Hopkins, 2008). Systematic comparison across all analyzed protein targets revealed a subset of compounds capable of engaging two or more disease-associated proteins (Figure 9). The presence of such multi-target interaction profiles suggests potential polypharmacological properties, which may be particularly advantageous in the context of diabetic nephropathy, a multifactorial disease characterized by the convergence of inflammatory, metabolic, and fibrotic signaling pathways.

By simultaneously modulating multiple targets, these ligands may offer enhanced therapeutic efficacy or improved pathway-level regulation compared to single-target compounds. The network-based identification of multi-target ligands therefore provides a rational framework for prioritizing candidates for further in silico validation and experimental evaluation, including binding mode analysis, pathway enrichment, and functional assays.

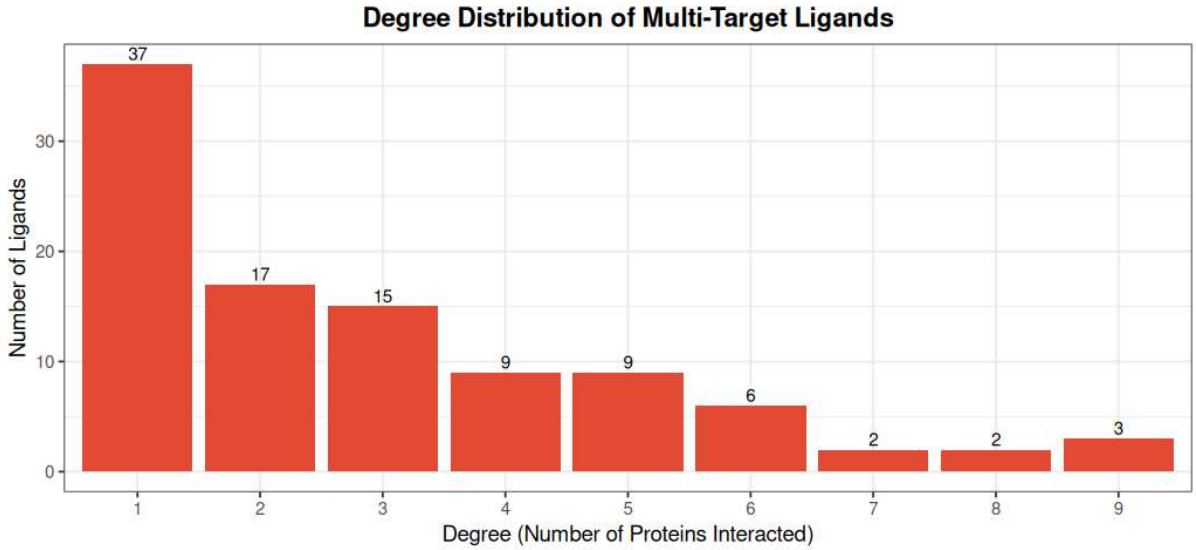


Figure 9: Degree distribution of multi-target ligands, representing the frequency of ligands based on their degree (number of connected protein targets; 1–9), thereby revealing the extent of ligand polypharmacology

Degree centrality analysis of the protein–ligand interaction network identified a subset of ligands exhibiting pronounced multi-target binding behaviour, demonstrating polypharmacological potential consistent with network pharmacology approaches (Singh et al., 2026b) with individual compounds interacting with six to nine distinct protein targets (Table 3). Such high-degree

ligands represent potential polypharmacological candidates, capable of simultaneously engaging multiple disease-relevant proteins.

Among these, LTS0003161, LTS0144363, and LTS0182645 demonstrated the highest network connectivity (degree = 9), indicating extensive interaction profiles across the analyzed target set. The elevated degree values of these ligands suggest a capacity to modulate multiple signaling pathways implicated in diabetic nephropathy, including inflammatory, immune, and metabolic processes. Consequently, these compounds emerge as prioritized candidates for further in-depth computational characterization and experimental validation.

Table 3: Top 10 Multitarget ligand with their mean binding energy

S.N.	Ligand ID	No. of Proteins	Interacting Proteins	Mean Binding Energy (kcal/mol)	Best Binding Energy (kcal/mol)
1	LTS0003161 (Ergostane-steroids)	9	5GJI, 3BH8, 5LWM, 6ZCJ, 7SIY, 4OM7, 1UZK, 5QU2, 6NEW	−9.98	−12.09
2	LTS0144363 (Ergostane-steroids)	9	5GJI, 3 BH8, 4P5M, 5GJH, 6ZCJ, 1RJB, 7P8X, 5QU2, 6NEW	−10.21	−12.01
3	LTS0182645 (Ergostane-steroids)	9	5P9J, 4P5M, 1KQ6, 8YWA, 2OQ1, 7P8X, 6NEW, 3RY4, 2IIM	−9.70	−11.64
4	LTS0185704 (Ergostane-steroids)	8	5GJH, 2WNV, 1E8Y, 5LWM, 4YJR, 4HKC, 1UZK, 6NMW	−10.01	−11.39
5	LTS0153798 (Ergostane-steroids)	8	5H0B, 6QZH, 4P5M, 8YWA, 1UGN, 2OQ1, 4HKC, 5QU2	−9.82	−10.83
6	LTS0056127 (Ergostane-steroids)	7	5GJI, 5GJH, 5P9J, 6ZCJ, 6QZH, 1RJB, 6NMW	−10.44	−11.92
7	LTS0168926 (Ergostane-steroids)	7	1E8Y, 5H0B, 1RJB, 3RY4, 2OQ1, 6NEW, 6NMW	−9.61	−11.86
8	LTS0062707 (Ergostane-steroids)	6	8JOT, 5GJH, 5LWM, 1RJB, 4OM7, 7P8X	−10.24	−11.00
9	LTS0069095 (Ergostane-steroids)	6	5P9J, 5H0B, 6QZH, 7SIY, 1RJB, 2OQ1	−10.05	−10.58
10	LTS0010381 (Ergostane-steroids)	6	5LWM, 4P5M, 1KQ6, 4HKC, 5QU2, 6NEW	−9.67	−10.37

Functional annotation of the protein targets interacting with the prioritized ligands revealed significant enrichment in pathways related to insulin signaling, inflammatory responses, oxidative stress, mitochondrial dysfunction, and extracellular matrix (ECM) remodeling—all of which represent core pathological processes underlying diabetic nephropathy. These functional categories collectively reflect the multifactorial nature of disease progression, encompassing metabolic dysregulation, chronic inflammation, cellular stress responses, and fibrotic remodeling within the diabetic kidney.

The convergence of ligand interactions on proteins involved in these pathways suggests that the identified compounds may exert broad, pathway-level modulatory effects, rather than acting on isolated molecular targets. Such multi-pathway engagement is consistent with the complex pathophysiology of diabetic nephropathy and supports the relevance of a polypharmacological strategy aimed at simultaneously addressing interconnected disease mechanisms (Singh et al., 2026b).

The consistently strong predicted binding affinities of these ligands across multiple disease-relevant protein targets underscore their potential as multi-target therapeutic candidates (Katara, 2020). Such interaction profiles suggest a capacity for simultaneous modulation of interconnected pathogenic pathways, which may be particularly advantageous in the treatment of diabetic nephropathy, a disorder driven by complex crosstalk between metabolic, inflammatory, oxidative, and fibrotic processes. While these findings are based on in silico predictions, they provide a compelling rationale for prioritizing these compounds for further validation, including detailed interaction analyses, molecular dynamics simulations, and experimental assays.

DISCUSSION

This integrative transcriptomics and network-guided investigation provides a systems-level framework for understanding the molecular architecture of diabetic nephropathy (DN). Recently, various studies have successfully investigated complex diseases and predicted therapeutic targets and compounds using transcriptomics-based integrative approaches (Szklarczyk et al., 2025; Wen et al., 2025). Considering the robustness of these methods, the present study employed an integrative framework to deduce DN-associated genes, their regulatory networks, drug targets, and therapeutic agents.

The current study utilized RNA-seq datasets and performed differential gene expression (DEG) analysis, identifying 1,484 upregulated and 591 downregulated genes. Our analysis focused on upregulated DEGs to prioritize targets for inhibitory screening, as these genes often represent the

hyperactive pathological cascades driven by metabolic stress (Odarenko et al., 2025). The stringent K-means clustering ($k=75$) resolved transcriptionally coordinated modules, revealing that larger clusters harbored a greater number of annotated metabolic drug targets. This suggests that highly populated transcriptional modules are not random gene aggregates but represent biologically active pathways enriched in therapeutically actionable genes. This trend supports the concept that metabolic dysregulation in DN is mediated through interconnected gene networks rather than isolated molecular events (He et al., 2025). Assessment of network robustness through clustering coefficient analysis identified 14 highly cohesive clusters (coefficients ≥ 0.7), providing a rational systems-biology framework for prioritizing biologically meaningful therapeutic modules (Zhou et al., 2020).

The resulting protein–protein interaction (PPI) network (Figure 4) revealed densely interconnected subnetworks enriched for immune activation, inflammatory signaling, and extracellular matrix (ECM) organization—processes widely recognized as central drivers of DN pathogenesis (Tuttle et al., 2014; Reidy et al., 2014). The emergence of hub genes such as *CD4, CCR7, CSF1R, FBN1, ITGA4, LCK, SYK,* and *TLR6* highlights immune cell–kidney crosstalk and ECM–receptor signaling as the primary integrative axes of disease progression. Given their high centrality, these hubs are well-positioned to propagate pathological signals, making them attractive candidates for therapeutic intervention (Barabási et al., 2011).

Regulatory network reconstruction identified transcription factors (e.g., *NFKB1, RELA, SP1, STAT3*) and miRNAs (e.g., *hsa-let-7a-5p, hsa-miR-21-5p*) as candidate master regulators of these hubs. The frequent occurrence of feed-forward loop motifs (Figure 5) suggests that DN-associated transcriptional states are reinforced by multilayered regulatory circuits, which may explain the limited efficacy of single-pathway therapeutic interventions (Shalgi et al., 2007). Furthermore, the identified miRNAs, which mediate post-transcriptional gene regulation, represent promising candidates for miRNA-based therapeutic intervention, enabling the concurrent modulation of multiple disease-relevant targets—a systems-level advantage over conventional single-gene targeting (Katara et al., 2010; Yang et al., 2023).

Functional enrichment analysis corroborated these findings, with KEGG pathway analysis emphasizing immune signaling and phagocytic pathways. While some pathways appeared annotated under infectious diseases (e.g., COVID-19, tuberculosis), these likely reflect shared inflammatory and immune-response genes rather than direct infection (Figure 6). Collectively, the enrichment profile suggests that DN progression involves a coordinated response of immune activation, oxidative stress, and fibrotic remodeling (Wada & Makino, 2013).

Our molecular docking analysis provided a systematic in silico framework to interrogate ligand–target interactions across 29 prioritized protein targets. The processing of 18,502 docking simulations enabled a comprehensive assessment of the polypharmacological potential of natural products. The construction of a protein–ligand interaction network highlighted compounds capable of binding multiple DN-associated targets, which is advantageous for a multifactorial disorder characterized by overlapping metabolic, inflammatory, and fibrotic pathways (Kabir & Muth, 2022). Notably, ergostane-type steroids (e.g., *LTS0003161, LTS0144363, LTS0182645*) emerged as the most prominent multi-target ligands, displaying favorable binding energies (approaching −12 kcal/mol) and interaction profiles that align with the pathogenic axes identified in our network analysis. These findings suggest that the same network properties that complicate DN treatment—namely the convergence of metabolic stress and immune activation—also create an opportunity for rational multi-target therapeutic strategies.

Limitations of the Study

While our integrative approach provides a robust framework, it is subject to several limitations. First, our analysis focused on upregulated DEGs to prioritize targets for inhibitory screening; however, we acknowledge that this approach excludes potentially protective genes whose downregulation may also contribute to DN pathology. Second, as this study relies on bulk tissue biopsy data, the enrichment of immune-related pathways reflects both intrinsic renal responses and the infiltration of immune cells, rather than purely disease-specific expression. Third, the docking analyses are in silico predictions; while we have prioritized ligands with strong theoretical binding affinities, these results represent hypothesis generation. The identified candidates require subsequent experimental validation—including binding affinity assays, molecular dynamics simulations, and in vivo efficacy studies—to confirm biological activity, pharmacokinetics, and safety profiles prior to any clinical translation.

CONCLUSION:

In conclusion, this study presents an integrated systems biology framework combining transcriptomics, network analysis, and structure-based virtual screening to elucidate the molecular complexity of diabetic nephropathy. By anchoring dysregulated gene modules to curated metabolic drug targets, the analysis demonstrates that DN progression is driven by coordinated immune, inflammatory, and extracellular matrix–associated networks rather than isolated molecular alterations. The identification of highly connected hub genes and multilayer regulatory circuits, i.e., Gene–TF–miRNA networks, underscores the central role of immune–

kidney cross talk and persistent transcriptional regulation in shaping disease progression. Importantly, integration of network topology with molecular docking highlights ergostane-type steroids from natural compounds as promising multi-target ligands. These ligands are capable of simultaneously engaging key proteins across major pathogenic axes, including insulin signaling, inflammation, oxidative stress, and fibrotic remodeling. Collectively, this work provides a resource of disease modules, regulatory drivers, and candidate compounds, laying a foundation for experimental validation and the rational development of multi-target therapeutics for diabetic nephropathy. As the study is based on publicly available datasets and in silico analyses, experimental validation in cellular, animal, and clinical models remains essential to confirm the biological relevance and therapeutic potential of the identified targets and multi-target ligands.

Acknowledgement: ARK acknowledges research fellowship support from the Department of Biotechnology (DBT-GoI) under grant no. DBT/2023-24/UoA/2306. MSK acknowledges research fellowship support from the University Grants Commission (UGC-GoI).

Declaration:

- **Competing Interests:** The authors declare that they have no conflict of interest.
- **Ethics, Consent to Participate, and Consent to Publish declarations**: Not applicable.
- **Financial Support:** No financial support was received for the conduct of this research.
- **Data Availability Statement:** All data generated or analysed during this study are included in this published article [and its supplementary information files].

References:

Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data. Available online at: *http://www.bioinformatics.babraham.ac.uk/projects/fastqc*

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., & Sherlock, G. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nature genetics*, *25*(1), 25–29. https://doi.org/10.1038/75556

Aswal, S., Kumar, A., Chauhan, A., Semwal, R. B., Kumar, A., & Semwal, D. K. (2020). A Molecular Approach on the Protective Effects of Mangiferin Against Diabetes and Diabetes-related Complications. *Current diabetes reviews*, *16*(7), 690–698. https://doi.org/10.2174/1573399815666191004112023

Barabási, A. L., Gulbahce, N., & Loscalzo, J. (2011). Network medicine: a network-based approach to human disease. *Nature reviews. Genetics*, *12*(1), 56–68. https://doi.org/10.1038/nrg2918

Bartel, D. P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, *116*(2), 281–297. https://doi.org/10.1016/s0092-8674(04)00045-5

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., & Bourne, P. E. (2000). The Protein Data Bank. *Nucleic Acids Research, 28*(1), 235–242. https://doi.org/10.1093/nar/28.1.235

BIOVIA, Dassault Systèmes. (2024). *Discovery Studio Visualizer (Version 24.1)*

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics (Oxford, England)*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Cai, F. H., Qian, F. C., Li, B. L., Li, L. D., Liao, B. H., Yu, Z. M., Fang, Q. L., Li, Y. Y., Dong, F. J., Zhou, L. W., Li, C., Wang, Q. Y., & Liu, J. H. (2025). DiabetesOmic: A comprehensive multi-omics diabetes database. *Computational and structural biotechnology journal*, *27*, 2147–2154. https://doi.org/10.1016/j.csbj.2025.05.008

Chang, L., Zhou, G., Soufan, O., & Xia, J. (2020). miRNet 2.0: network-based visual analytics for miRNA functional analysis and systems biology. *Nucleic acids research*, *48*(W1), W244–W251. https://doi.org/10.1093/nar/gkaa467

Chin, C. H., Chen, S. H., Wu, H. H., Ho, C. W., Ko, M. T., & Lin, C. Y. (2014). cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC systems biology*, *8 Suppl 4*(Suppl 4), S11. https://doi.org/10.1186/1752-0509-8-S4-S11

Cotas, J., Lomartire, S., Pereira, L., Valado, A., Marques, J. C., & Gonçalves, A. M. M. (2024). Seaweeds as Nutraceutical Elements and Drugs for Diabetes Mellitus: Future Perspectives. *Marine drugs*, *22*(4), 168. https://doi.org/10.3390/md22040168

Dong, F., Zheng, L., & Yang, G. (2024). Construction of a TF-miRNA-mRNA regulatory network for diabetic nephropathy. Archivos Españoles de Urología, 77(1), 104–112. https://doi.org/10.56434/j.arch.esp.urol.20247701.14

Eberhardt, J., Santos-Martins, D., Tillack, A. F., & Forli, S. (2021). AutoDock Vina 1.2.0: New docking methods, expanded force field, and Python bindings. *Journal of Chemical Information and Modeling*, *61*(8), 3891–3898. https://doi.org/10.1021/acs.jcim.1c00203

Ertl, P., Roggo, S., & Schuffenhauer, A. (2008). Natural product-likeness score and its application for prioritization of compound libraries. *Journal of chemical information and modeling*, *48*(1), 68–74. https://doi.org/10.1021/ci700286x

Fan, Y., Siklenka, K., Arora, S. K., Ribeiro, P., Kimmins, S., & Xia, J. (2016). miRNet - dissecting miRNA-target interactions and functional associations through network-based visual analysis. *Nucleic acids research*, *44*(W1), W135–W141. https://doi.org/10.1093/nar/gkw288

Fan, Y., Yi, Z., D'Agati, V. D., Sun, Z., Zhong, F., Zhang, W., Wen, J., Zhou, T., Li, Z., He, L., Zhang, Q., Lee, K., He, J. C., & Wang, N. (2019). Comparison of Kidney Transcriptomic Profiles of Early and Advanced Diabetic Nephropathy Reveals Potential New Mechanisms for Disease Progression. *Diabetes*, *68*(12), 2301–2314. https://doi.org/10.2337/db19-0204

Fan, Y., Yi, Z., D'Agati, V. D., Sun, Z., Zhong, F., Zhang, W., Wen, J., Zhou, T., Li, Z., He, L., Zhang, Q., Lee, K., He, J. C., & Wang, N. (2020). Erratum. Comparison of Kidney

Transcriptomic Profiles of Early and Advanced Diabetic Nephropathy Reveals Potential New Mechanisms for Disease Progression. Diabetes 2019;68:2301-2314. *Diabetes*, *69*(4), 797. https://doi.org/10.2337/db20-er04b

Forli Lab. (2022). Meeko: A Python package for preparing small molecules for AutoDock (Version 0.4.0) [Software]. Scripps Research. https://github.com/forlilab/Meeko

Gaonkar, V. P., & Hullatti, K. (2020). Indian Traditional medicinal plants as a source of potent Anti-diabetic agents: A Review. *Journal of diabetes and metabolic disorders*, *19*(2), 1895–1908. https://doi.org/10.1007/s40200-020-00628-8

Gong, J., Fang, X., Yao, X., Xue, F., Qi, G., Chen, B., & Sun, H. (2025). Molecular mechanisms and multi-target therapeutic strategies of diabetic nephropathy: from pathogenesis to precision interventions. *Frontiers in pharmacology*, *16*, 1729941. https://doi.org/10.3389/fphar.2025.1729941

He, S., Chen, Y. W., Ye, J., Wang, Y., Chen, Q. K., & Liu, S. Y. (2025). Investigating the metabolic reprogramming mechanisms in diabetic nephropathy: a comprehensive analysis using bioinformatics and machine learning. *Frontiers in cell and developmental biology*, *13*, 1630708. https://doi.org/10.3389/fcell.2025.1630708

Hopkins, A. L. (2008). Network pharmacology: the next paradigm in drug discovery. *Nature chemical biology*, *4*(11), 682–690. https://doi.org/10.1038/nchembio.118

Hu, Q., Jiang, L., Yan, Q., Zeng, J., Ma, X., & Zhao, Y. (2023). A natural products solution to diabetic nephropathy therapy. *Pharmacology & therapeutics*, *241*, 108314. https://doi.org/10.1016/j.pharmthera.2022.108314

Jonas, S., & Izaurralde, E. (2015). Towards a molecular understanding of microRNA-mediated gene silencing. *Nature reviews. Genetics*, *16*(7), 421–433. https://doi.org/10.1038/nrg3965

Kabir, A., & Muth, A. (2022). Polypharmacology: The science of multi-targeting molecules. *Pharmacological research*, *176*, 106055. https://doi.org/10.1016/j.phrs.2021.106055

Katara, P. (2020). Computational Approaches for Drug Target Identification. In: Singh, D.B. (eds) Computer-Aided Drug Design. Springer, Singapore. https://doi.org/10.1007/978-981-15-6815-2_8

Katara, P., Gautam, B., Kuntal, H., & Sharma, V. (2010). Prediction of miRNA targets, affected proteins and their homologs in Glycine max. *Bioinformation*, *5*(4), 162–165. https://doi.org/10.6026/97320630005162

Koutrouli, M., Karatzas, E., Paez-Espino, D., & Pavlopoulos, G. A. (2020). A Guide to Conquer the Biological Network Era Using Graph Theory. *Frontiers in bioengineering and biotechnology*, *8*, 34. https://doi.org/10.3389/fbioe.2020.00034

Li, L., Jin, W., Kim, J., Bae, G., Yang, S. H., Cho, B., Han, S. H., Lee, J., Kim, D., Kim, D. K., Lim, C. S., Hong, B. H., & Lee, J. P. (2025). Graphene Quantum Dots as Antifibrotic Therapy for Kidney Disease. *ACS applied bio materials*, *8*(2), 980–991. https://doi.org/10.1021/acsabm.4c01053

Liao, Y., Smyth, G. K., & Shi, W. (2019). The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic acids research*, *47*(8), e47. https://doi.org/10.1093/nar/gkz114

Lipinski, C. A., Lombardo, F., Dominy, B. W., & Feeney, P. J. (2001). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced drug delivery reviews*, *46*(1-3), 3–26. https://doi.org/10.1016/s0169-409x(00)00129-0

Liu, F., Peng, A., Zhu, X., & Wang, G. (2025). Differential expression and correlation analysis of whole transcriptome for type 2 diabetes mellitus. *Frontiers in endocrinology*, *16*, 1541261. https://doi.org/10.3389/fendo.2025.1541261

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology*, *15*(12), 550. https://doi.org/10.1186/s13059-014-0550-8

Ma, Z., Hou, Q., Yang, R., Shi, J., Kan, S., Zheng, C., Zhu, X., Liu, Z., & Jiang, S. (2025). Transcriptome-wide association study revealed novel causal genes of renal-biopsy proven diabetic nephropathy. *Genome medicine*, *18*(1), 6. https://doi.org/10.1186/s13073-025-01590-x

Martínez-López, J. A., Lindqvist, A., Lopez-Pascual, A., Harder, A., Chen, P., Ngara, M., Shcherbina, L., Siffo, S., Cowan, E., Baira, S. M., Kryvokhyzha, D., Karagiannopoulos, A., Chriett, S., Skene, N. G., Prasad, R. B., Lancien, M., Johnson, P. F., Eliasson, P., Eliasson, L., Louvet, C., … Wierup, N. (2025). Single-cell mRNA-regulation analysis reveals cell type-specific mechanisms of type 2 diabetes. *Nature communications*, *16*(1), 9475. https://doi.org/10.1038/s41467-025-65060-z

Mezzano, S., Aros, C., Droguett, A., Burgos, M. E., Ardiles, L., Flores, C., Schneider, H., Ruiz-Ortega, M., & Egido, J. (2004). NF-kappaB activation and overexpression of regulated genes in human diabetic nephropathy. Nephrology, dialysis, transplantation : official publication of the European Dialysis and Transplant Association - European Renal Association, 19(10), 2505–2512. https://doi.org/10.1093/ndt/gfh207

Mima, A., Nomura, A., & Yasuzawa, T. (2025). Update on the pathophysiology and treatment of diabetic kidney disease: a narrative review. *Expert review of clinical immunology*, *21*(7), 921–928. https://doi.org/10.1080/1744666X.2025.2521086

Mora, C., & Navarro, J. F. (2004). Inflammation and pathogenesis of diabetic nephropathy. *Metabolism: clinical and experimental*, *53*(2), 265–267. https://doi.org/10.1016/j.metabol.2003.11.005

Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., & Olson, A. J. (2009). AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *Journal of computational chemistry*, *30*(16), 2785–2791. https://doi.org/10.1002/jcc.21256

Navarro-González, J. F., Mora-Fernández, C., Muros de Fuentes, M., & García-Pérez, J. (2011). Inflammatory molecules and pathways in the pathogenesis of diabetic nephropathy. *Nature reviews. Nephrology*, *7*(6), 327–340. https://doi.org/10.1038/nrneph.2011.51

Nogales, C., Mamdouh, Z. M., List, M., Kiel, C., Casas, A. I., & Schmidt, H. H. H. W. (2022). Network pharmacology: curing causal mechanisms instead of treating symptoms. *Trends in pharmacological sciences*, *43*(2), 136–150. https://doi.org/10.1016/j.tips.2021.11.004

Odarenko, K. V., Zenkova, M. A., & Markov, A. V. (2025). Transcriptomic-Driven Drug Repurposing Reveals SP600125 as a Promising Drug Candidate for the Treatment of Glial-Mesenchymal Transition in Glioblastoma. *International journal of molecular sciences*, *26*(19), 9772. https://doi.org/10.3390/ijms26199772

Pérez-Morales, R. E., Del Pino, M. D., Valdivielso, J. M., Ortiz, A., Mora-Fernández, C., & Navarro-González, J. F. (2019). Inflammation in Diabetic Kidney Disease. *Nephron*, *143*(1), 12–16. https://doi.org/10.1159/000493278

Pun, F. W., Podolskiy, D., Izumchenko, E., Mortlock, A., Oprea, T. I., Scheibye-Knudsen, M., Fortney, K., Morgen, E., Ren, F., & Zhavoronkov, A. (2026). Target identification and assessment in the era of AI. *Nature reviews. Drug discovery*, 10.1038/s41573-026-01412-8. Advance online publication. https://doi.org/10.1038/s41573-026-01412-8

Rana, A., Samtiya, M., Dhewa, T., Mishra, V., & Aluko, R. E. (2022). Health benefits of polyphenols: A concise review. *Journal of food biochemistry*, *46*(10), e14264. https://doi.org/10.1111/jfbc.14264

Reddy, M. A., & Natarajan, R. (2013). Role of epigenetic mechanisms in the vascular complications of diabetes. *Sub-cellular biochemistry*, *61*, 435–454. https://doi.org/10.1007/978-94-007-4525-4_19

Reidy, K., Kang, H. M., Hostetter, T., & Susztak, K. (2014). Molecular mechanisms of diabetic kidney disease. *The Journal of clinical investigation*, *124*(6), 2333–2340. https://doi.org/10.1172/JCI72271

Reutens, A. T., & Atkins, R. C. (2011). Epidemiology of diabetic nephropathy. *Contributions to nephrology*, *170*, 1–7. https://doi.org/10.1159/000324934

Ruan, J., Dean, A. K., & Zhang, W. (2010). A general co-expression network-based approach to gene expression analysis: comparison and applications. *BMC systems biology*, *4*, 8. https://doi.org/10.1186/1752-0509-4-8

Rupaimoole, R., & Slack, F. J. (2017). MicroRNA therapeutics: towards a new era for the management of cancer and other diseases. *Nature reviews. Drug discovery*, *16*(3), 203–222. https://doi.org/10.1038/nrd.2016.246

Samsu N. (2021). Diabetic Nephropathy: Challenges in Pathogenesis, Diagnosis, and Treatment. *BioMed research international*, *2021*, 1497449. https://doi.org/10.1155/2021/1497449

Schrödinger, LLC. (2015). The PyMOL molecular graphics system, version 1.8. (https://pymol.org)

Setten, R. L., Rossi, J. J., & Han, S. P. (2019). The current state and future directions of RNAi-based therapeutics. *Nature reviews. Drug discovery*, *18*(6), 421–446. https://doi.org/10.1038/s41573-019-0017-4

Shalgi, R., Lieber, D., Oren, M., & Pilpel, Y. (2007). Global and local architecture of the mammalian microRNA-transcription factor regulatory network. *PLoS computational biology*, *3*(7), e131. https://doi.org/10.1371/journal.pcbi.0030131

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., & Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research*, *13*(11), 2498–2504. https://doi.org/10.1101/gr.1239303

Singh, N., Kori, A. R., & Katara, P. (2026a). scRNA-seq based gene and cell dynamics behind the formation of the non-caseating granulomas. *Computational biology and chemistry*, 124 (Pt 1), 109070. https://doi.org/10.1016/j.compbiolchem.2026.109070

Singh, N., Kori, A. R., & Katara, P. (2026b). Uncovering multi-target natural inhibitors for hypertension through network pharmacology and structure-based screening. *Network Biology, 16*(2), 31–48.

Stuart, J. M., Segal, E., Koller, D., & Kim, S. K. (2003). A gene-coexpression network for global discovery of conserved genetic modules. *Science (New York, N.Y.)*, *302*(5643), 249–255. https://doi.org/10.1126/science.1087447

Szklarczyk, D., Nastou, K., Koutrouli, M., Kirsch, R., Mehryary, F., Hachilif, R., Hu, D., Peluso, M. E., Huang, Q., Fang, T., Doncheva, N. T., Pyysalo, S., Bork, P., Jensen, L. J., & von Mering, C. (2025). The STRING database in 2025: protein networks with directionality of regulation. *Nucleic acids research*, *53*(D1), D730–D737.https://doi.org/10.1093/nar/gkae1113

Tang, S. C. W., & Yiu, W. H. (2020). Innate immunity in diabetic kidney disease. *Nature reviews. Nephrology*, *16*(4), 206–222. https://doi.org/10.1038/s41581-019-0234-4

Tao, K., Shao, L., Chang, H., Chen, X., Xia, H., Wu, R., Wang, S., & Liao, H. (2025). Integrative Transcriptomic and Network Pharmacology Analysis Reveals Key Targets and Mechanisms of *Moschus* (musk) Against Viral Respiratory Tract Infections. *Pharmaceuticals (Basel, Switzerland)*, *18*(8), 1136. https://doi.org/10.3390/ph18081136

Treiber, T., Treiber, N., & Meister, G. (2019). Regulation of microRNA biogenesis and its crosstalk with other cellular pathways. *Nature reviews. Molecular cell biology*, *20*(1), 5–20. https://doi.org/10.1038/s41580-018-0059-1

Trott, O., & Olson, A. J. (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of computational chemistry*, *31*(2), 455–461. https://doi.org/10.1002/jcc.21334

Tuttle, K. R., Bakris, G. L., Bilous, R. W., Chiang, J. L., de Boer, I. H., Goldstein-Fuchs, J., Hirsch, I. B., Kalantar-Zadeh, K., Narva, A. S., Navaneethan, S. D., Neumiller, J. J., Patel, U. D., Ratner, R. E., Whaley-Connell, A. T., & Molitch, M. E. (2014). Diabetic kidney disease: A report from an ADA consensus conference. *Diabetes Care*, *37*(10), 2864–2883. https://doi.org/10.2337/dc14-1296

Verma, Y., Yadav, A., & Katara, P. (2020). Mining of cancer core-genes and their protein interactome using expression profiling based PPI network approach. *Gene Reports, 18*, 100583. https://doi.org/10.1016/j.genrep.2019.100583

Wada, J., & Makino, H. (2013). Inflammation and the pathogenesis of diabetic nephropathy. *Clinical science (London, England : 1979)*, *124*(3), 139–152. https://doi.org/10.1042/CS20120198

Wang, K., Wu, Y. G., Su, J., Zhang, J. J., Zhang, P., & Qi, X. M. (2012). Total glucosides of paeony regulates JAK2/STAT3 activation and macrophage proliferation in diabetic rat kidneys. *The American journal of Chinese medicine*, *40*(3), 521–536. https://doi.org/10.1142/S0192415X12500401

Wang, Y., Chen, Y., Liang, X., Zhu, L., & Wen, X. (2024). Network pharmacology and transcriptomics explore the therapeutic effects of Ermiao Wan categorized formulas for diabetes in mice. *Scientific reports*, *14*(1), 27014. https://doi.org/10.1038/s41598-024-78364-9

Wen, W., Li, R., Zhuo, P., Deng, C., & Zhang, D. (2025). Integrated network toxicology, machine learning and molecular docking reveal the mechanism of benzopyrene-induced periodontitis. *BMC Pharmacology and Toxicology*, *26*(1), 118. https://doi.org/10.1186/s40360-025-00961-9

Wishart, D. S., Feunang, Y. D., Guo, A. C., Lo, E. J., Marcu, A., Grant, J. R., Sajed, T., Johnson, D., Li, C., Sayeeda, Z., Assempour, N., Iynkkaran, I., Liu, Y., Maciejewski, A., Gale, N., Wilson, A., Chin, L., Cummings, R., Le, D., Pon, A., … Wilson, M. (2018). DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic acids research*, *46*(D1), D1074–D1082. https://doi.org/10.1093/nar/gkx1037

World Health Organization. (2024). *Diabetes: Key facts*. https://www.who.int/news-room/fact-sheets/detail/diabetes

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. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation (Cambridge (Mass.))*, *2*(3), 100141. https://doi.org/10.1016/j.xinn.2021.100141

Yamagishi, S., & Matsui, T. (2010). Advanced glycation end products, oxidative stress and diabetic nephropathy. *Oxidative medicine and cellular longevity*, *3*(2), 101–108. https://doi.org/10.4161/oxim.3.2.11148

Yang, H., Liu, Y., Chen, L., Zhao, J., Guo, M., Zhao, X., Wen, Z., He, Z., Chen, C., & Xu, L. (2023). MiRNA-Based Therapies for Lung Cancer: Opportunities and Challenges? *Biomolecules*, *13*(6), 877. https://doi.org/10.3390/biom13060877

Zhang, L., Jiang, L., Xu, R., Zhang, X., Zhang, B., & Yue, R. (2025). Epidemiological research on diabetic nephropathy at global, regional, and national levels from 1990 to 2021: an analysis derived from the global burden of disease 2021 study. *Frontiers in endocrinology*, *16*, 1647064. https://doi.org/10.3389/fendo.2025.1647064

Zhou, L. T., Qiu, S., Lv, L. L., Li, Z. L., Liu, H., Tang, R. N., Ma, K. L., & Liu, B. C. (2018). Integrative Bioinformatics Analysis Provides Insight into the Molecular Mechanisms of Chronic Kidney Disease. *Kidney & blood pressure research*, *43*(2), 568–581. https://doi.org/10.1159/000488830

Zhou, Y., Hou, Y., Shen, J., Huang, Y., Martin, W., & Cheng, F. (2020). Network-based drug repurposing for novel coronavirus 2019-nCoV/SARS-CoV-2. *Cell discovery*, *6*, 14. https://doi.org/10.1038/s41421-020-0153-3

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryTableS1.csv
- SupplementaryTableS6.csv
- SupplementaryTableS3.xlsx
- SupplementaryTableS2.csv
- SupplementaryTableS7.csv
- SupplementaryTableS8.csv
- SupplementaryTableS9.csv
- SupplementaryTableS4.csv
- SupplementaryTableS10.csv
- SupplementaryTableS14.csv
- SupplementaryTableS13.csv
- SupplementaryTableS11.csv
- SupplementaryTableS12.csv
- SupplementaryTableS15.csv
- supplementarytableslist.docx
- SupplementaryTableS5.csv