Identification of hub genes associated with accumulation of extracellular matrix as biomarkers of diabetic nephropathy

Song-Tao Feng
doctoral student
Yue-Ming Gao
doctoral student
Di Yin
doctoral student
Lin-Li Lv
colleague
Yi Wen
colleague
Zuo-Lin Li
colleague
Bin Wang
colleague
Min Wu
colleague
Bi-Cheng Liu (iubc64@163.com)
Institute of Nephrology, Zhong Da Hospital, Southeast University School of Medicine
https://orcid.org/0000-0002-4584-1853

Research

Keywords: Diabetic nephropathies, Gene regulatory networks, Extracellular matrix, Biomarkers

Posted Date: July 6th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-40090/v1

Abstract

Background

Diabetic nephropathy (DN) remains a major cause of end stage renal disease (ESRD). The development of novel biomarkers and early diagnosis of DN are of great clinical importance. The goal of this study was to identify hub genes with diagnostic potential for DN by weighted gene co-expression network analysis (WGCNA).

Methods

Gene Expression Omnibus (GEO) database was searched for microarray data including distinct types of chronic kidney diseases (CKD). Gene co-expression network was constructed and modules specific for DN were identified by WGCNA. Gene Ontology (GO) analysis was performed and the hub genes were screened out within the selected gene modules. Furthermore, receiver operating characteristic (ROC) curves were generated to evaluate the diagnostic values of hub genes. In addition, an external validation was performed in an independent dataset.

Results

Dataset GSE99339 was selected and a total of 179 microdissected glomeruli samples were analyzed, including DN, normal control and 7 groups of other glomerular diseases. 23 modules of the total 10947 genes were grouped by WGCNA and a module was specifically correlated with DN (r = 0.54, 9e-15). GO analysis showed that module genes were mainly enriched in the accumulation of extracellular matrix (ECM). LUM, ELN, FBLN1, MMP2, FBLN5 and FMOD were identified as hub genes. Furthermore, levels of hub genes were the highest in DN compared to other groups, which could differentially diagnose DN (AUC, 0.67 ~ 0.95). External verification showed hub genes were higher in DN group and were negatively correlated with eGFR.

Conclusions

By using WGCNA approach, we identified 6 hub genes, LUM, ELN, FBLN1, MMP2, FBLN5 and FMOD, related to ECM accumulation and were specific for DN. These genes may represent potential candidate diagnostic biomarkers of DN.

Background

Diabetic nephropathy (DN), a severe microvascular complication of diabetes mellitus (DM), represents a crucial cause of chronic kidney disease (CKD) which frequently leads to end stage renal disease (ESRD). DN develops in about 30% of patients with type1 DM (T1DM) and 40% of patients with type 2 DM (T2DM)

[1]. Clinically, DN is characterized by progressive renal damage reflected by persistent albuminuria, decline of glomerular filtration rate (GFR), hypertension, and excess mortality due to ESRD or cardiovascular complications. A report from American Diabetes Association (ADA) indicated that DN was the most common cause of ESRD in many countries in the world including the United States, Japan and Europe. Patients with DN accounted for 25–45% of patients enrolled in ESRD programs [2]. Taking into account the high incidence and heavy health and public burden of ESRD, the early diagnosis and justified management of DN are of great clinical importance.

The clinical diagnosis of DN is based on estimated GFR (eGFR) and albuminuria, combined with the clinical characteristics of DM, such as the course of disease and the existence of diabetic retinopathy [3]. DN is defined as persistently urinary albumin to creatinine ratio (ACR) $\geq$ 30 mg/g and/or sustained decline in eGFR < 60 ml/min/1.73 m^2 [1, 4]. For patients with atypical clinical manifestations of DM and renal involvement, renal biopsy remains the “golden standard” of final diagnosis. Morphological changes include thickness of glomerular and tubular basement membranes, mesangial expansion, and typical glomerulosclerosis with nodular mesangial lesions (Kimmelstiel-Wilson nodule). However, the elevated level of albuminuria and the reduction of eGFR are not exclusive to DN, which are challenged when patients with DM develop CKD with atypical clinical features. The application of renal biopsy subjects to great restrictions because of its strict indications and the risk of internal bleeding [5]. There is an increasing need for development of novel biomarkers of DN to allow for more accurate diagnosis.

Technologies of microarray and RNA sequencing have made great progress in the past two decades. As an algorithm of systems biology, weighted gene co-expression network analysis (WGCNA) enables the identification of clusters of biologically related genes and it can associate the gene clusters (often called gene modules) with specific phenotypes or clinical traits, such as the pathological classification of disease. WGCNA has emerged as a powerful approach to understand mechanisms and identify potential biomarkers or therapeutic targets [6–8]. In view of its advantages, analyzing the whole transcriptional characteristics is of great significance for understanding pathogenesis and discovery of specific biomarkers for glomerular diseases. In 2012, by comparing the differentially expressed genes (DEGs) between DN and normal subjects, Wang et al [9] reported gene modules and several transcriptional factors related to the development of DN. To our best knowledge, WGCNA has so far not applied to identify specific genes that distinguish DN from other pathological types of CKD.

In the present study, we analyzed whole-transcriptome gene expression data of microdissected glomeruli samples from various types of CKD including DN. WGCNA was used to group global genes into modules in an unsupervised manner and the modules specific for DN were identified. Further bioinformatics tools, like gene ontology (GO) analysis and protein-protein interaction (PPI) network, allowed the identification of the most connected and central genes (called hub genes). The main goal of this study was to expand our understanding of the mechanism of DN, unravelling specific novel candidate biomarkers that can differentiate DN from other types of glomerular diseases.

Materials And Methods

Microarray data processing

The mRNA expression matrix of GSE99339 [10] was downloaded from the Gene Expression Omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/geo/). This dataset was performed on the GPL19109 platform (Affymetrix Human Genome U133 Plus 2.0 Array). mRNA expression data was preprocessed and analyzed in R statistical environment (version 3.6.1). Background correction and normalization were performed by “normalize-BetweenArrays” function in the R package “limma” (version 3.6) in Bioconductor project (version 3.1) (http://www.-bioconductor.org/). After the initial filtering, the expression data of 10,947 probes were yielded. Then, the array annotation R package “hgu133plus2.db” (version 3.6) was used to map the microarray probes to the official gene symbol.

Construction of gene co-expression networks and identification of gene modules related to DN.

Gene co-expression network analysis was performed using the R package “WGCNA” (version 1.69), a co-expression network analysis approach proposed by Steve Horvath in 2005 [11]. The details of algorithm referred to the WGCNA manual (https://cran.r-project.org/web/packages/WGCNA/WGCNA.pdf). Briefly, Pearson correlation coefficient was calculated for all the genes. Then, a soft power threshold was used to transform the correlation matrix into a weighted adjacency matrix. Next, the adjacency matrix was transformed into a topological overlap matrix (TOM) and 1-TOM was calculated. Modules identification was accomplished with the method of dynamic tree cut by hierarchically clustering genes taking 1-TOM as the distance measure with a minimum module size cut-off of 30 genes. Highly similar modules were recognized by clustering and then merged together with a height cut-off of 0.25. Then, the module eigengenes (MEs) were regarded as gene expression profiles that best characterize the overall levels of modules. MEs were then analyzed for correlations with clinical traits of subjects, that was, the pathological diagnosis of renal biopsy. The gene significance was defined as the correlation between gene and clinical traits [12]. The MEs of interests were selected by clinical traits of diagnosis of DN based on GS > 0.4 and with a threshold of *P*-value < 0.05.

Gene ontology (GO) enrichment analysis

To interpret the biological function of the interested modules, the visualization of GO enrichment terms of the selected module associated with DN was performed using R package “clusterProfiler” (version 3.11) [13] in the Bioconductor project. Adjust *P* value < 0.05 was considered as statistically significant.

Recognition of hub genes and further analysis

The online platform, STRING (http://string-db.org), was utilized to construct protein-protein interaction (PPI) networks within the interested modules. Those with the highest degrees, defined as the edges’ number incident to the nodes, were recognized as hub genes. “Cytoscape” software (version 3.7.2) was utilized to analyze the characteristics of each node of PPI and hub genes were identified by Maximal Clique Centrality (MCC) method in “cyto-hubba”, a plug-in of Cytoscape software. The log2 transformed expression levels of hub genes were compared in different groups of glomerular diseases. Receiver operator characteristic (ROC) curves and the areas under ROC (AUCs) were generated to evaluate the

efficacy of hub genes in the diagnosis of DN different from normal control and other pathological types of glomerular diseases.

External verification

Dataset GSE30528, a study consisted of mRNA profiles of 9 cases of glomeruli samples with DN and 13 healthy living donors, was downloaded from GEO database and used as external verification. The log2 transformed mRNA expression levels of hub genes were compared between DN and control group (healthy living donors). ROC curves were used to evaluate the efficacy of hub genes in diagnosis of DN. The clinical data of GSE30528 was downloaded from “Nephroseq” online open-access platform (http://v5.nephroseq.org). Pearson correlation was performed between mRNA levels of hub genes and eGFR values.

Statistical analysis

Statistical analysis was performed in R and GraphPad Prism 7.0 (GraphPad Software, Inc.). One-way ANOVA analysis or Kruskal-Wallis test was performed for comparisons among groups. Dunnett’s or Dunn’s test was used for multiple comparisons. Unpaired *t* test or Mann-Whitney test was used to evaluate statistical significance between the values of patients with DN and control group. The ROC curves were established and AUCs were calculated to evaluate the efficiency of hub genes in diagnosing DN. Pearson correlation coefficient was performed to assess the correlations between mRNA levels and eGFR. All tests were two tailed, with a value of $P < 0.05$ considered as statistically significant.

Results

Gene expression data

Dataset GSE99339 contains mRNA expression data in a total of 187 microdissected glomeruli samples from renal biopsy tissues, including 14 cases of DN, 22 cases of focal segmental glomerular sclerosis (FSGS), 15 cases of hypertensive nephropathy (HT), 26 cases of IgA nephropathy (IgAN), 13 cases of minimal change disease (MCD), 21 cases of membranous glomerulonephritis (MGN), 23 cases of rapidly progressive glomerular nephritis (RPGN), 30 cases of lupus nephritis (LN), 14 cases of normal control of tumor nephrectomy (TN), 6 cases of FSGS&MCD and 3 cases of thin basement membrane nephropathy (TMD) respectively. Cases of FSGS&MCD were excluded for the reason of ambiguity of diagnosis. Cases of TMD were excluded since the number of cases was relatively small.

Gene modules relevant to DN was identified via WCGNA

To identify gene modules associated with pathological types of DN, WGCNA was applied based on the whole-transcriptome data of microdissected glomeruli samples. Power of 8 was selected to approximate scale-free topology (Fig. 1a). Modules were generated by dynamic tree cut and were merged with the number of genes less than 30 and the cutting heights of 0.25 (Fig. 1b, c). Then we got 23 modules for the whole transcriptome (Fig. 1d). MEs were calculated as the representative for modules. Pearson correlation coefficients were calculated between MEs and the clinical traits of pathological type. From the heatmap of

module-trait correlations, we identified that the saddle brown module was the highest correlated with DN (r = 0.54, P = 9e-5), assigning 64 genes (Additional file 1: Table S1). The correlations with other groups of glomerular diseases were relatively weak (Fig. 1d).

GO terms were mainly enriched in ECM accumulation and 6 genes were identified as hub genes

In order to annotate the biological functions of genes in saddle brown module, the GO enrichment analysis was performed. GO analysis showed genes were mainly enriched in the component and biological process of extracellular matrix (ECM) (Fig. 2a). In addition, we constructed a PPI network (Fig. 2b) of the module genes. Hub genes with the highest degrees of connectivity in the PPI network were applied to MCC method (Additional file 1: Table S2). Genes of LUM, ELN, FBLN1, MMP2, FBLN5 and FMOD were recognized as hub genes with the top 6 MCC score (Fig. 2c, Table 1).

Table 1
The description of hub genes

Gene symbol	Official full name	Description[a]
LUM	lumican	A member of the SLRP family. Distributed in interstitial collagenous matrices.
ELN	elastin	Components of elastic fibers. Comprise part of the extracellular matrix.
FBLN1	fibulin 1	Secreted glycoprotein that becomes incorporated into a fibrillar extracellular matrix.
MMP2	matrix metallopeptidase 2	A member of the MMP gene family. Capable of cleaving components of the extracellular matrix.
FBLN5	fibulin 5	A secreted, extracellular matrix protein. Induced in atherosclerotic lesions, notably in endothelial cells.
FMOD	fibromodulin	A member of the SLRP family. May play a role in the assembly of extracellular matrix and TGF-beta activities.
a: The description of hub genes was extracted from the NCBI database (https://www.ncbi.nlm.nih.gov/).		
SLRP: small leucine-rich proteoglycans.		

The expression level and diagnostic capacity of hub genes in DN group

As shown in in Fig. 3A-F, the median log2 transformed mRNA levels of LUM, ELN, FBLN1, MMP2, FBLN5 and FMOD in DN group were the highest among all the groups, although the increase of some genes is not statistically significant. In detail, LUM, ELN, and FMOD were statistically highest in DN group. For FBLN1, the level in DN was statistically higher than that of FSGS, MGN and LN groups. For MMP2, the level in DN group was statistically higher than that of TN, MCD, HT, IgAN, FSGS and MGN groups. For FBLN5, except TN and groups, the expression level in DN was statistically higher than that of other groups.

These 6 hub genes showed relatively high AUCs levels for differential diagnosis of DN from other glomerular diseases. The AUCs of LUM, FBLN1, MMP2, FBLN5 and FMOD in the differential diagnosis of DN from other groups were higher than 0.7 (0.7078 ~ 0.9524). (Table 2). FMOD showed the best efficiency of discrimination (AUC, 0.8864 ~ 0.9381).

Table 2
The area under the ROC curves of hub genes in differential diagnosing DN

Gene symbol	Area under the ROC curve							
	DN/TN	DN/MCD	DN/HT	DN/IgAN	DN/FSGS	DN/MGN	DN/LN	DN/RPGN
LUM	0.8776	0.8352	0.8048	0.7885	0.7565	0.9048	0.7881	0.7019
ELN	0.8878	0.8516	0.8286	0.783	0.9091	0.8401	0.8905	0.7547
FBLN1	0.6786	0.7363	0.7619	0.739	0.8052	0.8435	0.8024	0.736
MMP2	0.9388	0.9286	0.919	0.8187	0.8279	0.8878	0.7881	0.764
FBLN5	0.7857	0.7143	0.781	0.7582	0.7078	0.8946	0.7905	0.7609
FMOD	0.9082	0.8956	0.9238	0.9258	0.8864	0.9524	0.9381	0.8727

ROC: Receiver operator characteristic; DN: Diabetic nephropathy; TN: Tumor nephrectomy; MCD: Minimal change disease; HT: Hypertensive nephropathy; IgAN: IgA nephropathy; FSGS: Focal segmental glomerular sclerosis; MGN: Membranous glomerulonephritis; RPGN: Rapidly progressive glomerular nephritis; LN: Lupus nephritis.

External validation

Another independent microarray dataset GSE30528, including 9 cases of glomeruli samples with DN and 13 cases of normal control, was used for external validation. The clinical data, including age, sex, serum creatinine and eGFR was downloaded from Nephroseq platform. The log2 transformed mRNA levels of hub genes and values of eGFR of subjects in GSE30528 were downloaded from Nephroseq platform (Additional file 1: Table S3). As shown in Fig. 4a-f. Except ELN, the expression levels of other 5 hub genes in DN group were statistically higher than that of normal control group ($P < 0.01$). The ROC curves showed that except ELN (AUC, 0.624), other 5 genes had relatively high AUCs in the diagnosis of DN (AUC, 0.778 ~ 0.992) (Fig. 5a-f). Furthermore, Pearson correlation analysis showed the expression levels of 6 hub genes were negatively correlated with eGFR (Fig. 6a-f). Except ELN (r=-0.176, $P = 0.533$), the correlations of other 5 hub genes with eGFR were statistically significant (r = -0.487 ~-0.658, $P < 0.05$).

Discussion

In the present study, microarray data from 179 pathologically confirmed renal biopsy samples was analyzed, covering a variety of primary or secondary glomerular diseases. By defining the pathological types as clinical traits and analyzing the correlation with gene modules, a cluster of genes with high expression specificity in DN was generated. WGCNA was performed in an independent study of large number of samples of multiple kidney diseases, avoiding the heterogeneity among different studies

caused by a comprehensive analysis of multiple GEO datasets. Different from the analysis of DEGs [14–16], the genome-wide transcriptional co-expression network analysis has provided an unbiased description of gene expression network in DN. A gene module specific for DN was identified reflecting the unique pathogenesis of DN. GO analysis revealed genes assigned by the module were mainly enriched in the components and biological process of ECM, which were consistent with the pathological characteristics of DN. Next, we identified 6 hub genes including LUM, ELN, FBLN1, MMP2, FBLN5 and FMOD in the specific module and found that DN showed the highest expression levels. Furthermore, LUM, FBLN1, MMP2, FBLN5 and FMOD showed efficient diagnostic capacity of DN. In addition, these findings were proved by external validation.

The hallmark of the pathogenesis of DN is an increased ECM accumulation which causes thickness of the glomerular and tubular basement membranes, followed by mesangial expansion, glomerular sclerosis, and tubulointerstitial fibrosis. The main components of ECM in DN include elevated expression of collagen I, collagen III, collagen IV (α1 and α2 chains), collagen V, collagen VI, fibronectin, laminin, and small leucine-rich proteoglycans (SLRP) [17–18]. The major physiologic regulators of ECM degeneration in glomeruli were matrix metalloproteinases (MMPs), mainly MMP-2 and MMP-9 [19]. Interestingly, among the 6 hub genes identified in this study, LUM and FMOD encode members of SLR, lumican and fibromodulin and MMP-2 was critical in the ECM accumulation of DN.

Lumican and fibromodulin, members of class II SLRP family [20] encoded by LUM and FMOD, were basically expressed in kidney [21–22]. In DN, lumican and fibromodulin were mainly localized in areas of sclerosis scars and become progressively more evident with the extent of fibrosis in glomeruli and tubulointerstitium [23]. As soluble forms of lumican and fibromodulin could be released into body fluid [23], Ahn et al [24] reported lumican was identified belong to 13 up-regulated glycoproteins in DN plasma by multi-lectin affinity chromatography. Fibulin-1, a secreted glycoprotein encoded by FBLN1, was highly expressed in the capillary loop of DN glomeruli [25]. The plasma level in glomerulonephritis, DN and obstructive nephropathy were elevated compared with normal subjects [26]. In addition, Scholze, et al [27] found that plasma fibulin-1 was related to cardiovascular risk in CKD patients, especially in patients of CKD and DM. MMP-2 degrades the components of ECM, including col-IV, fibronectin, aggrecan, laminin, elastin, gelatin, and non-matrix substrates. In patients with DN, MMP-2 was increased in kidney samples and urinary MMP-2 was elevated in patients with albuminuria and established renal injury [28, 29]. Elastin, encoded by ELN, was found accumulated in DN kidney samples and involved in the deposition of ECM in DN [30]. In the present study, its diagnostic value was not so satisfactory as that of the other 5 hub genes.

DN is usually a clinical diagnosis according to albuminuria and/or reduced eGFR in the absence of signs of other causes of renal disease. As renal biopsy is not indispensable to DN diagnosis, a proportion of non-diabetic renal disease indicated by renal biopsy is among patients clinically diagnosed with DN, especially MGN and IgAN [31, 32]. The development of DN diagnostic biomarkers in plasma or urine will contribute to accurate clinical diagnosis. As mentioned above, lumican, fibromodulin, fibulin and MMP-2 can exist in blood or urine in soluble form, and the levels were elevated in patients with DKD or DM compared with

normal subjects. It is of great significance to verify whether these hub genes have specific diagnostic values for DN in clinical utility.

Limitations

The main limitation of this study lies in the lack of established clinical data of GSE99339. In addition, it is a data mining research based on bioinformatic analysis. There is still a need of validation cohort and direct evidences that hub genes can be used as biomarkers of DN, although the external validation in an independent dataset in our study supported our findings.

Conclusions

In conclusion, through WGCNA and other bioinformatics tools, we identified a gene module specifically associated with DN, which enriched in component and biological process of ECM accumulation. LUM, ELN, FBLN1, MMP2, FBLN5 and FMOD were identified as hub genes within the gene module, which showed efficiency in differential diagnosis of DN. These findings provided a novel approach and candidate potential biomarkers for the diagnosis of DN, which can stimulate researchers engaged in clinical verification and help to develop novel biomarkers of DN.

Abbreviations

DN: Diabetic nephropathy; DM:Diabetes mellitus; CKD:Chronic kidney disease; ESRD:End stage renal disease; GFR:Glomerular filtration rate; ADA:American Diabetes Association; ACR:Albumin to creatinine ratio; WGCNA:Weighted gene co-expression network analysis; GO:Gene ontology; PPI:Protein-protein interaction; GEO:Gene Expression Omnibus; TOM:Topological overlap matrix; ME:Module eigengenes; MCC:Maximal Clique Centrality; ROC:Receiver operator characteristic; AUC:area under ROC curve; ECM:Extracellular matrix; DEGs:Differential expressed genes; SLRP:Small leucine-rich proteoglycan; MMP:Matrix metalloproteinases; TN:Tumor nephrectomy; MCD:Minimal change disease; HT:Hypertensive nephropathy; IgAN:IgA nephropathy; FSGS:Focal segmental glomerular sclerosis; MGN:Membranous glomerulonephritis; RPGN:Rapidly progressive glomerular nephritis; LN:Lupus nephritis.

Declarations

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Funding

This study was supported by grants from the National Key Research and Development Program (2018YFC1314000) and the National Natural Science Foundation of China (81670696).

Authors' Contributions

Author's contributions: BW and BCL designed the study. DY searched the microarray and clinical data of datasets. STF and YMG performed the WGCNA and bioinformatic analysis of this study. YW and ZLL prepared the tables and figures. STF and HL drafted manuscript. LLL and BCL edited and revised manuscript. BCL approved final version of manuscript. All authors read and approved the final manuscript.

Acknowledgement

The authors thank Dr. Shi Zhang for his kind suggestions about methods of WGCNA algorithm.

References

1. Alicic RZ, Rooney MT, Tuttle KR. Diabetic kidney disease: challenges, progress, and possibilities. Clin J Am Soc Nephro. 2017;12(12):2032–45.
2. Tuttle KR, Bakris GL, Bilous RW, Chiang JL, de Boer IH, Goldstein-Fuchs J, et al. Diabetic kidney disease: a report from an ADA consensus conference. Diabetes Care. 2014;37(10):2864-83.
3. Cefalu WT, Berg EG, Saraco M, Petersen MP, Uelmen S, Robinson S, et al. Microvascular complications and foot care: standards of medical care in diabetes-2019. Diabetes Care. 2019;42:124-S38.
4. National Kidney Foundation. KDOQI clinical practice guideline for diabetes and CKD: 2012 update. Am J Kidney Dis. 2012;60(5):850–86.
5. Atwell TD, Spanbauer JC, McMenomy BP, Stockland AH, Hesley GK, Schleck CD, et al. The timing and presentation of major hemorrhage after 18,947 image-guided percutaneous biopsies. Am J Roentgenol. 2015;205(1):190–5.
6. Luo YP, Coskun V, Liang AB, Yu JH, Cheng LM, Ge WH, et al. Single-cell transcriptome analyses reveal signals to activate dormant neural stem cells. Cell. 2015;161(5):1175–86.

7. Duan HM, Ge WH, Zhang AF, Xi Y, Chen ZH, Luo DD, et al. Transcriptome analyses reveal molecular mechanisms underlying functional recovery after spinal cord injury. P Natl Acad Sci USA. 2015;112(43):13360–5.
8. Beckerman P, Qiu CX, Park J, Ledo N, Ko YA, Park ASD, et al. Human kidney tubule-specific gene expression based dissection of chronic kidney disease traits. Ebiomedicine. 2017;24:267–76.
9. Tang W, Gao Y, Li Y, Shi G. Gene networks implicated in diabetic kidney disease. Eur Rev Med Pharmaco. 2012;16(14):1967–73.
10. Shved N, Warsow G, Eichinger F, Hoogewijs D, Brandt S, Wild P, et al. Transcriptome-based network analysis reveals renal cell type-specific dysregulation of hypoxia-associated transcripts. Sci Rep-Uk. 2017;7.
11. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. Bmc Bioinformatics. 2008;9.
12. Bao C, Lu YK, Chen JS, Chen DN, Lou WY, Ding BS, et al. Exploring specific prognostic biomarkers in triple-negative breast cancer. Cell Death Dis. 2019;10.
13. Yu GC, Wang LG, Han YY, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. Omics. 2012;16(5):284–7.
14. Geng XD, Wang WW, Feng Z, Liu R, Cheng XL, Shen WJ, et al. Identification of key genes and pathways in diabetic nephropathy by bioinformatics analysis. J Diabetes Invest. 2019;10(4):972–84.
15. Zeng MR, Liu JL, Yang WX, Zhang SM, Liu FY, Dong Z, et al. Identification of key biomarkers in diabetic nephropathy via bioinformatic analysis. J Cell Biochem. 2019;120(5):8676–88.
16. Fu FM, Wei XL, Liu JB, Mi NR. Bioinformatic analysis of specific genes in diabetic nephropathy. Ren Fail. 2015;37(7):1219–24.
17. Mason RM, Wahab NA. Extracellular matrix metabolism in diabetic nephropathy. J Am Soc Nephrol. 2003;14(5):1358–73.
18. Kolset SO, Reinholt FP, Jenssen T. Diabetic nephropathy and extracellular matrix. J Histochem Cytochem. 2012;60(12):976–86.
19. Dimas GG, Didangelos TP, Grekas DM. Matrix gelatinases in atherosclerosis and diabetic nephropathy: progress and challenges. Curr Vasc Pharmacol. 2017;15(6):557–65.
20. Schaefer L. Small leucine-rich proteoglycans in kidney disease. J Am Soc Nephrol. 2011;22(7):1200–7.
21. Schaefer L, Grone HJ, Raslik I, Robenek H, Ugorcakova J, Budny S, et al. Small proteoglycans of normal adult human kidney: distinct expression patterns of decorin, biglycan, fibromodulin, and lumican. Kidney Int. 2000;58(4):1557–68.
22. Ying SX, Shiraishi A, Kao CWC, Converse RL, Funderburgh JL, Swiergiel J, et al. Characterization and expression of the mouse lumican gene. J Biol Chem. 1997;272(48):30306–13.
23. Schaefer L, Raslik I, Grone HJ, Schonherr E, Macakova K, Ugorcakova J, et al. Small proteoglycans in human diabetic nephropathy: Discrepancy between glomerular expression and protein accumulation of decorin, biglycan, lumican, and fibromodulin. Faseb J. 2001;15(3):559–61.

24. Ahn JM, Kim BG, Yu MH, Lee IK, Cho JY. Identification of diabetic nephropathy-selective proteins in human plasma by multi-lectin affinity chromatography and LC-MS/MS. Proteom Clin Appl. 2010;4(6–7):644–53.
25. Satoskar AA, Shapiro JP, Bott CN, Song HJ, Nadasdy GM, Brodsky SV, et al. Characterization of glomerular diseases using proteomic analysis of laser capture microdissected glomeruli. Modern Pathol. 2012;25(5):709–21.
26. Neiman M, Hedberg JJ, Donnes PR, Schuppe-Koistinen I, Hanschke S, Schindler R, et al. Plasma profiling reveals human fibulin-1 as candidate marker for renal impairment. J Proteome Res. 2011;10(11):4925–34.
27. Scholze A, Bladbjerg EM, Sidelmann JJ, Diederichsen ACP, Mickley H, Nybo M, et al. Plasma concentrations of extracellular matrix protein fibulin-1 are related to cardiovascular risk markers in chronic kidney disease and diabetes. Cardiovasc Diabetol. 2013;12.
28. Garcia-Fernandez N, Jacobs-Cacha C, Mora-Gutierrez JM, Vergara A, Orbe J, Soler MJ. Matrix metalloproteinases in diabetic kidney disease. J Clin Med. 2020;9(2).
29. Kim SS, Shin N, Bae SS, Lee MY, Rhee H, Kim IY, et al. Enhanced expression of two discrete isoforms of matrix metalloproteinase-2 in experimental and human diabetic nephropathy. Plos One. 2017;12(2).
30. Thongboonkerd V, Barati MT, McLeish KR, Benarafa C, Remold-O'Donnell E, Zheng SR, et al. Alterations in the renal elastin-elastase system in type 1 diabetic nephropathy identified by proteomic analysis. J Am Soc Nephrol. 2004;15(3):650–62.
31. Wang JL, Han QQ, Zhao LJ, Zhang JL, Wang YT, Wu YC, et al. Identification of clinical predictors of diabetic nephropathy and non-diabetic renal disease in Chinese patients with type 2 diabetes, with reference to disease course and outcome. Acta Diabetol. 2019;56(8):939–46.
32. Zhuo L, Zou GM, Li WG, Lu JH, Ren WW. Prevalence of diabetic nephropathy complicating non-diabetic renal disease among Chinese patients with type 2 diabetes mellitus. Eur J Med Res. 2013;18.

Figures

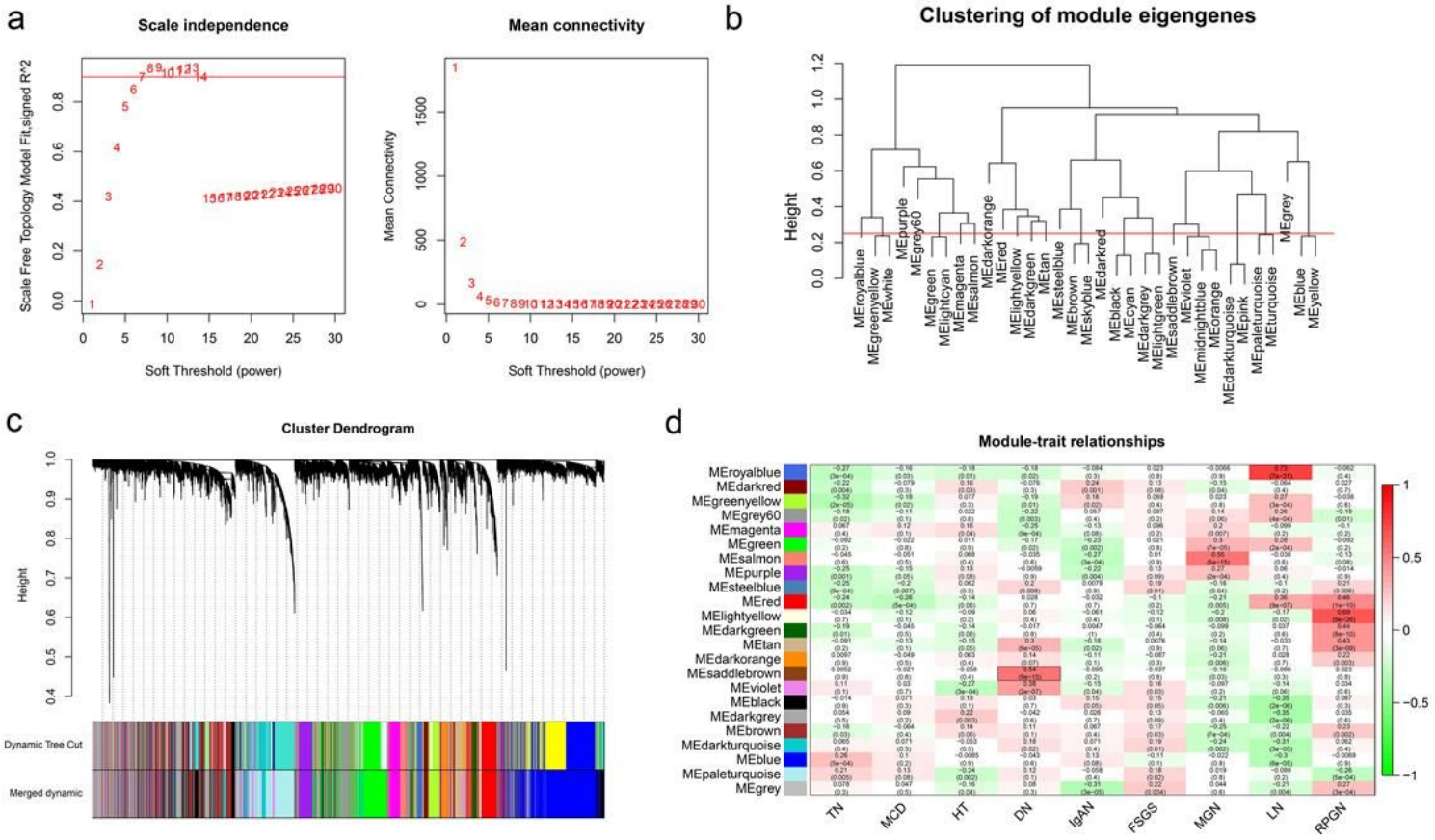

Figure 1

Identification of modules specially correlated with DN by WGCNA. a. Analysis of the scale-free fit index (left) and the mean connectivity (right) for various soft-thresholding powers. b. Clustering of module eigengenes. The red line indicates cut height (0.25). c. Dendrogram of all genes clustered based on a dissimilarity measure (1-TOM). d. Heatmap of the correlation between module eigengenes and pathological type. The saddle brown module was significantly correlated with DN. ME module eigengene, TN tumor nephrectomy, MCD minimal change disease, HT hypertensive nephropathy, DN diabetic nephropathy, IgAN IgA nephropathy, FSGS focal segmental glomerular sclerosis, MGN membranous glomerulonephritis, RPGN rapidly progressive glomerular nephritis, LN lupus nephritis

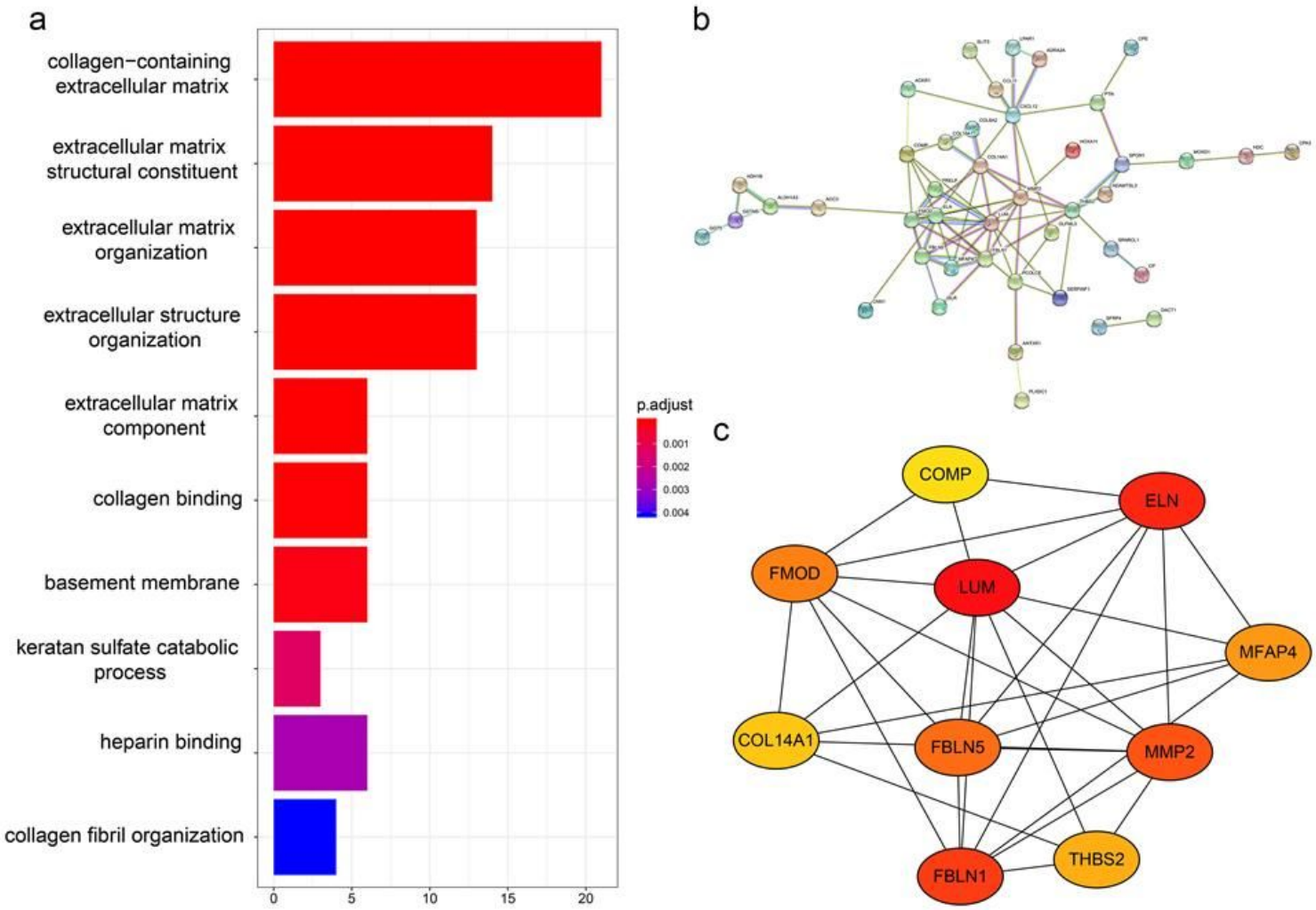

Figure 2

GO enrichment analysis and dentification of hub genes in the module genes highly correlated with DN. A. The top 10 GO terms enrichment analysis of genes assigned in saddle brown module. B. PPI network of genes assigned in saddle brown module. C. Hub genes of the PPI network of genes assigned in saddle brown module. GO gene ontology, PPI protein-protein network

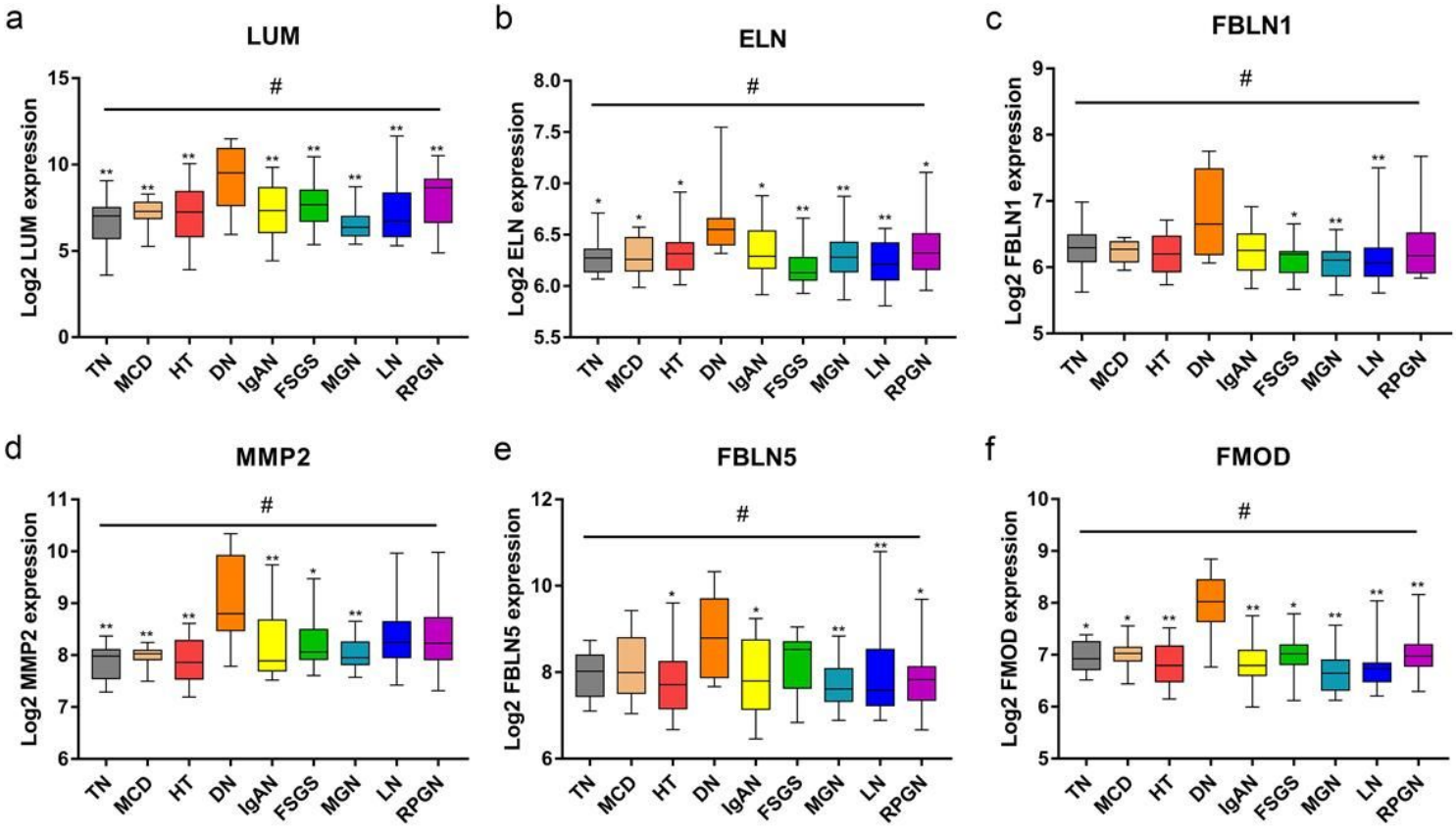

Figure 3

mRNA level of hub genes (a: LUM, b: ELN, c: FBLN1, d: MMP2, e: FBLN5, f: FMOD) in each group of kidney disease. TN tumor nephrectomy, MCD minimal change disease, HT hypertensive nephropathy, DN diabetic nephropathy, IgAN IgA nephropathy, FSGS focal segmental glomerular sclerosis, MGN membranous glomerulonephritis, LN lupus nephritis, RPGN rapidly progressive glomerular nephritis. # $P<0.05$, analyzed by one way-ANOVA analysis or Kruskal-Wallis test; * $P<0.05$, ** $P<0.01$, *** $P<0.001$, analyzed by multiple comparisons with DN group by Dunnett's or Dunn's test.

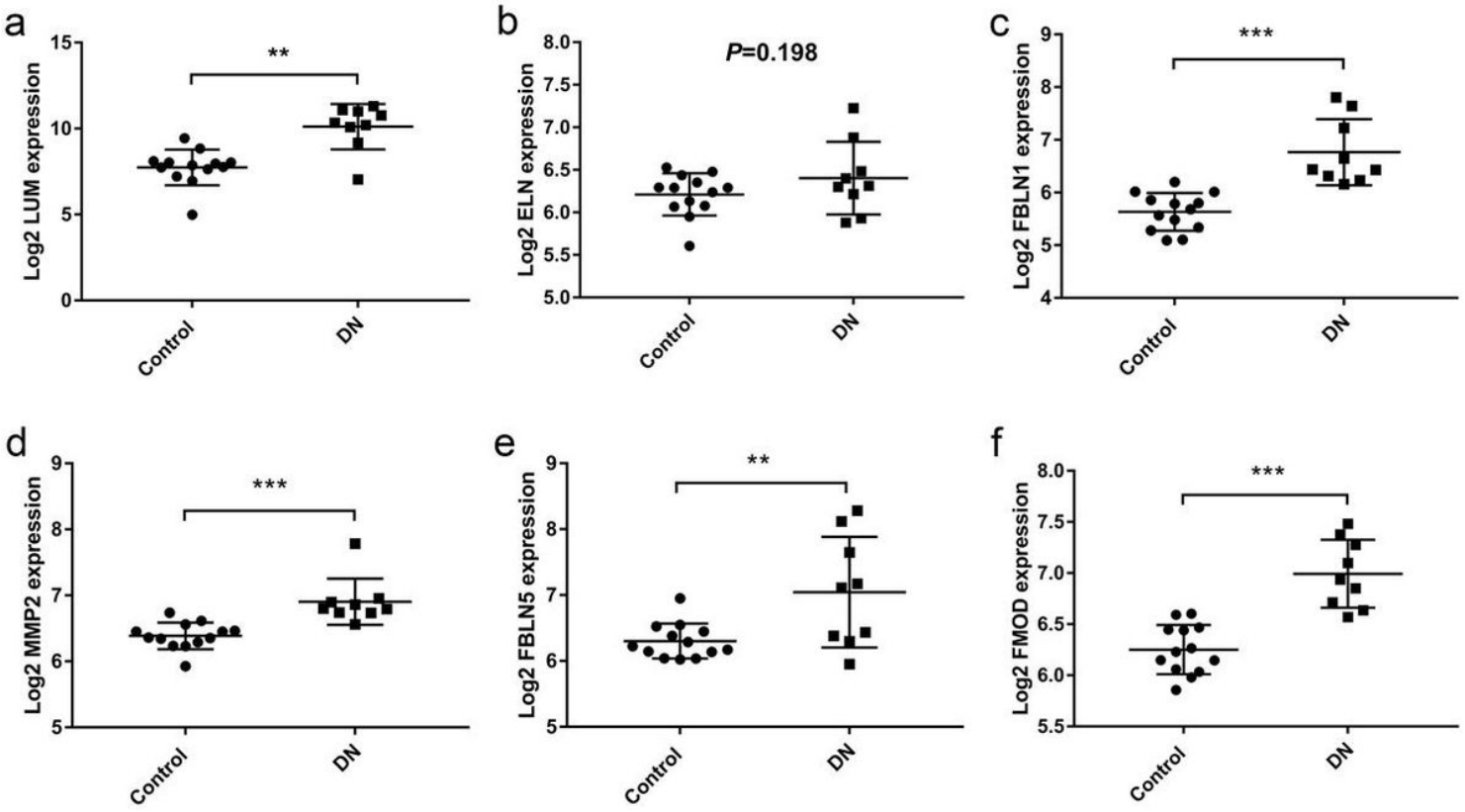

Figure 4

mRNA level of hub genes (a: LUM, b: ELN, c: FBLN1, d: MMP2, e: FBLN5, f: FMOD) in dataset GSE30528, as an external validation. DN diabetic nephropathy, control healthy individual donators. * P<0.05, ** P<0.01, *** P<0.001, analyzed by unpaired t-test or Mann-Whitney test.

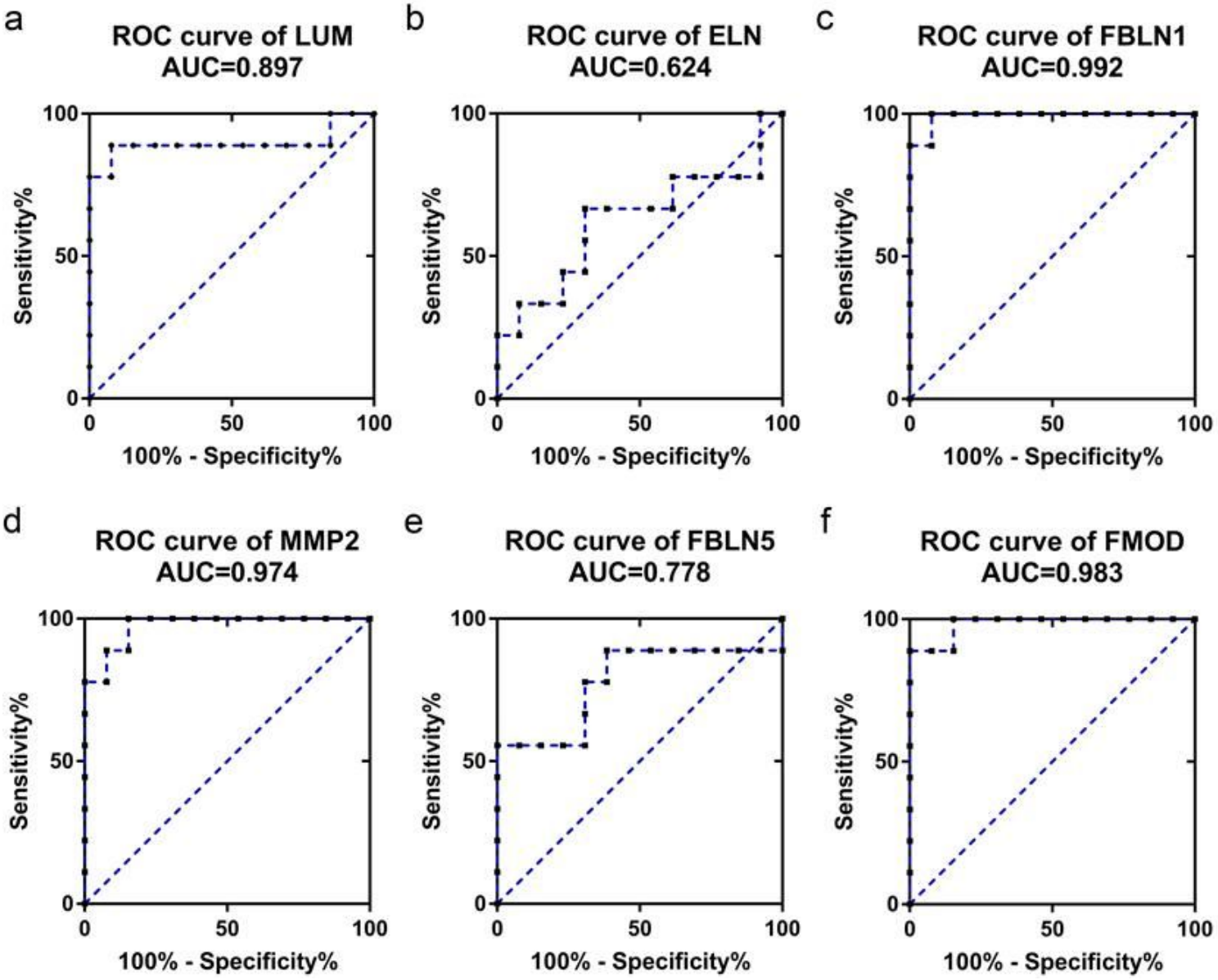

Figure 5

ROC curves of hub genes (a: LUM, b: ELN, c: FBLN1, d: MMP2, e: FBLN5, f: FMOD) in dataset GSE30528. The AUCs were calculated. A value of 0.5 is no better than that expected by chance (the null hypothesis). A value of 1.0 reflected a perfect indicator.

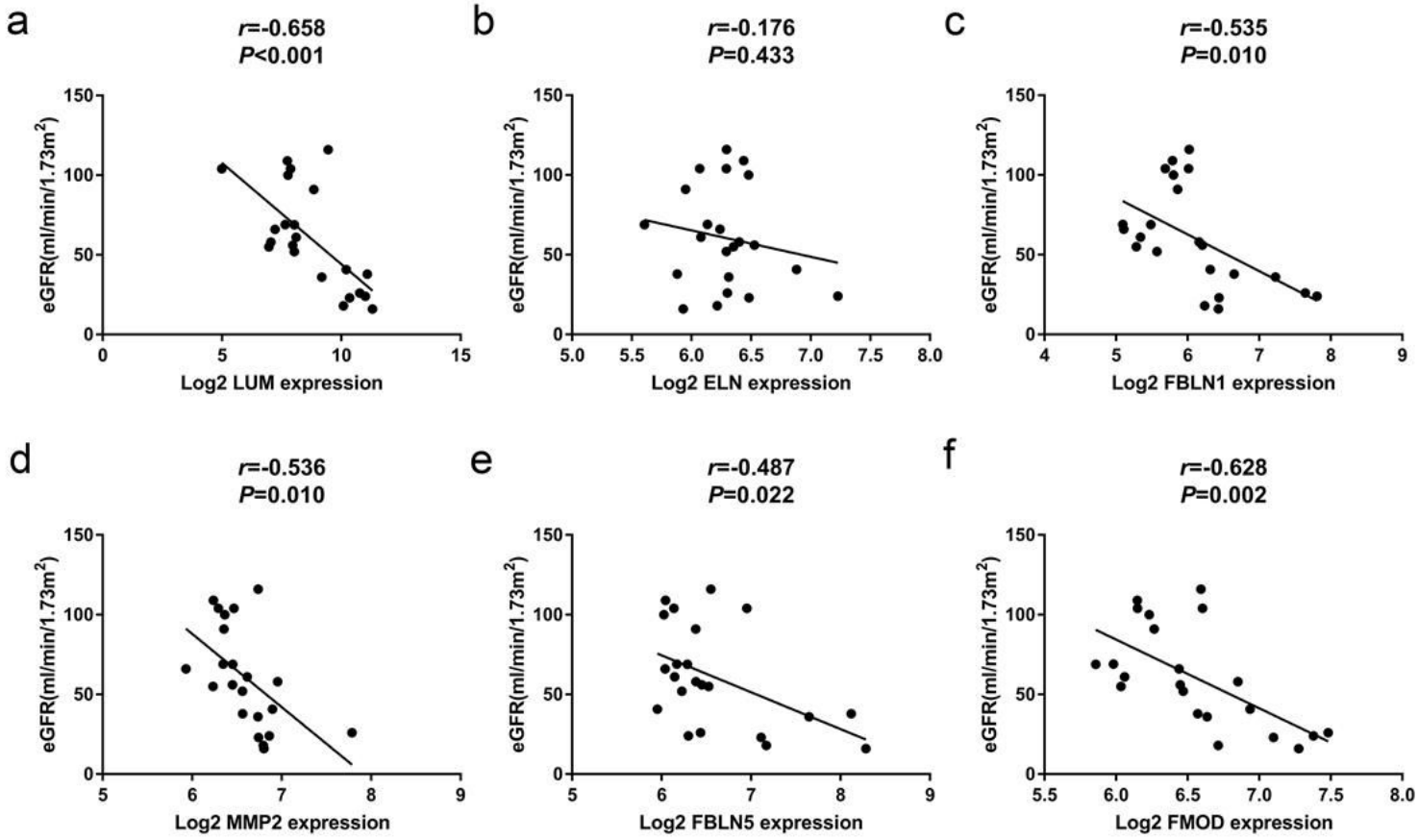

Figure 6

Pearson correlation of log2 transformed mRNA levels of hub genes (a: LUM, b: ELN, c: FBLN1, d: MMP2, e: FBLN5, f: FMOD) and eGFR in dataset GSE30528.

Supplementary Files

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

- TableS3.docx
- TableS2.docx
- TableS1.docx