Complement C3 as therapeutic target in diabetic nephropathy by bioinformatics analysis

ShuMei Tang
Xiangya Hospital Central South University https://orcid.org/0000-0003-3346-8614
XiuFen Wang
Department of Nephrology, Xiangya Hospital, Chang Sha, China
TianCi Deng
Department of Nephrology, Xiangya Hospital, Chang Sha, China
HuiPeng Ge
Department of Nephrology, Xiangya Hospital, Chang Sha, China
XiangCheng Xiao (1376785378@qq.com)
Department of Nephrology, Xiangya Hospital, Chang Sha, China

Research

Keywords: multi-omics, diabetic nephropathy, biomarkers, complement system, potential drugs

Posted Date: February 29th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-15517/v1

Abstract

Background: Various public platforms have contained extensive data for deeper bioinformatics analysis. The pathogenesis of diabetic nephropathy is always a hot topic and the underlying molecular events are not completely clear.

Methods: Differential expression analysis and weighted correlation network analysis (WGCNA) were applied to screen meaningful gene in GSE30529. The combined result was used for gene ontology (GO) annotation and KEGG pathway enrichment analysis. STRING was used for protein-protein interaction (PPI) network construction and Cytoscape software was used for hub gene identification. Gene methylation profile GSE121820 and miRNA profile GSE51674 were also acquired to perform multi-omics analysis. Clinical features were obtained form Nephroseq and potential drugs were identified by CMap.

Results: 345 genes were obtained form GSE30529 after differential expression analysis and WGCNA. GO analysis mainly included neutrophil activation, regulation of immune effector process, positive regulation of cytokine production and neutrophil mediated immunity. KEGG pathway analysis mostly included phagosome, complement and coagulation cascades, cell adhesion molecules and AGE-RAGE signaling pathway in diabetic complications. 16 down regulated miRNAs were obtained from GSE57674 to construct miRNA-mRNA network. Top 20 genes were discerned from PPI network and there were DNA methylation differences in 15 genes among them. Correlation analysis showed SYK, CXCL1, LYN, VWF, ANXA1, C3, HLA-E, RHOA, SERPING1, EGF and KNG1 may involved in DN. 10 small molecule compounds have been identified as potential therapeutic drugs.

Conclusion: We screened 11 target genes and suggested C3 may serve as a therapeutic target in diabetic nephropathy.

Background

It is estimated a total of 451 million people sufferd from diabetes by 2017 and the number is speculateed to 693 million by 2045 (1). As one of the most serious microvascular complication, diabetic nephropathy (DN) has been a major cause of end-stage renal disease (ESRD) in many countries. The congregation of advanced glycation end-products, oxidative stress and activation of proteinkinase C are the major contributing causes to the pathogenesis of DN. The new viewpoint holds that tubular injury plays an important and even initial role (2). Current treatment strategies for DN are aimed at controlling blood glucose and blood pressure levels and inhibition of RAS system to reduce albuminuria and delay the progression of DN (3). However, the effect is not entirely satisfactory for high incidence of DN-related ESRD. Therefore, there is a critical need to identify new therapeutic targets and improve clinical management.

High-throughput sequencing technology offers a effective method to study disease-related genes and provides promising medication goals in many fields, especially tumors (4). So far, several studies screened gene or miRNA involved in DN (5–9). Integrating these data could overcome the heterogeneity of researches and provide more accurate information. This study identified target genes whicn may improve the understanding of molecualr mechanisms of DN and provide a resource to build new hypotheses for further follow-up studies. We suggested complement system may serve as a therapeutic target in DN.

Methods

Data download

Expression profile GSE30529 (5) and miRNA profile GSE51674 (9) were downloaded by GEOquery package (10) in R software version 3.6.2. DNA methylation profile GSE121820_T2DN-CTL (unpublished) were downloaded from GEO database (https://www.ncbi.nlm.nih.gov/geo/). GSE30529 conducted by GPL571 covers 10 DN tubule samples and 12 control samples. GSE51674 conducted by GPL10656 includes 6 DN tissue samples and 4 control samples. GSE121820 conducted by GPL5082 contains 10 Type 2 DN blood samples and 10 control samples. The probe ID was converted into HUGO Gene Nomenclature Committee symbol (gene symbol).

Data processing

Differential expression analysis is the most commonly used method for data analysis. After checking the homogeneity of sample data in GSE30529, differentially expressed genes (DEGs) between DN samples and control samples were filtered with the criteria of |log2 fold change (FC)| greater than 1 and adjusted P-Value less than 0.05 by limma package (11). Similarly, differentially expressed miRNA (DERs) in GSE51674 were screened with |log2 FC| greater than 3 and adjusted P-Value less than 0.01.

Weighted correlation network analysis (WGCNA) allows biologically meaningful module information mining based on pairwise correlations between genes in high-throughput data using WGCNA package (12). WGCNA workflow consists of gene co-expression network construction, module identification, module relationship analysis and key drives gene recognition. The gene co-expression network was constructed with filtering principle that soft threshold is to make the network more consistent with a scale free topology. The module was identificated with criterion of module size 30 to 10000, merge cut height equal to 0.25 and verbose equal to 3. Highly related genes were obtained with threshold greater than 0.1 in Topological overlap matrix (TOM).

Target gene prediction of miRNA

TargetScan (13), miRWalK (14), miRBase (15) and miRTarBase (16) all allows for miRNA target prediction. The union set was used to miRNA-mRNA network construction.

Enrichment analysis and hub gene screening

Gene ontology (GO) annotation and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were performed with clusterProfiler package (17). The STRING database (18) (version 11.0, http://string-db.org/) is able to search interactions between candidate proteins based on laboratory data, other databases, text mining and predictive bioinformatics data. The protein-protein interaction (PPI) network was constituted by STRING database. Cytoscpe software was used to visualization and network analysis. In addition, CytoHubba, built-in tool in Cytoscape, allows 12 methods to explore important nodes in biological networks, such as Degree method (Deg), Maximum Neighborhood Component (MNC), Density of Maximum Neighborhood Component (DMNC), Maximal Clique Centrality (MCC), Closeness, EcCentricity, Radiality, BottleNeck, Stress, Betweenness, Edge Percolated Component (EPC) and ClusteringCofficient (19).

Clinical data analysis and drugs analysis

Nephroseq v5 analysis engine (http://v5.nephroseq.org) gives access to gene expression signatures and clinical features. The Pearson correlation analysis was performed between genes and GFR, proteinuria, and SCR level (5, 20). Unpaired Student t test was used to compare two groups. P-Value less than 0.05 was considered statistically significant. Insignificant results are not displayed.

Potential drugs were predicted with Connectivity Map (21), a online database that related disease, genes, and drugs based on similar or opposite gene expression signatures.

Results

Differential expression analysis of GSE30529

In order to determine whether data samples have batch effects or sample heterogeneity in GSE30529, gene expression levels of each sample were displayed in boxplot. And the results showed that they located approximately at the same level (Fig. 1A). In order to obtain the reliability of sample grouping, principal component analysis was performed in DN group and control group. The two main components respectively contributed 25.7% and 25.4% (Fig. 1B). Next, differential expression analysis was performed to acquire DEGs which may involved in DN with the criteria of |log2 FC| greater than 1 and adjusted P-Value less than 0.05 by limma package. As a result, 386 up-regulated DEGs and 71 down-regulated DEGs between DN samples and control samples were identified. The volcano map displayed the expression change and top 25 DEGs among them were shown in the hierarchical clustering heatmap (Fig. 1C and 1D).

Weighted correlation network analysis of GSE30529

WGCNA package is a set of functions used to calculate various weighted correlation analyses, which can be used for network construction, gene screening, gene module identification and topology feature calculation. Similarly, we performed sample cluster analysis first to learn sample similarity and grouping reliability, and result showed that there was 3 outliers (Fig. 2A). Three samples (GSM757025, GSM757027 and GSM757034) were removed. To construct a scale free network, the soft threshold power value was set as 10 with setting the soft threshold as 0.85 and mean connectivity threshold as 100 (Fig. 2B). Based on gene weighted expression correlation, hierarchical clustering analysis was carried out to obtain different gene modules, which were represented by branches of clustering tree and different colors. A total of 22 modules were found in network with module size identified as 30 to 10000 and merge cut hight identified as 0.25 (Fig. 2C). The 22 modules were divided into two clusters in general according to relationships between modules (Fig. 2D). In addition, weighted correlations of all genes (TOM) were displayed in heatmap plot (Fig. 2E). Fianllay, 3538 highly related genes were selected in TOM matrix with threshold greater than 0.1.

GO annotation and KEGG pathway analysis of combined result

DEG list and highly related gene list were combined to obtain more accurate targets (Fig. 3A). We obtained 345 genes uesd for GO annotation and KEGG pathway analysis to get bioinformatics annotation. Each gene is assigned at least three GO terms and each GO term is assigned at least five genes in chord chart. Top 12 GO terms were displayed and mainly included neutrophil activation, regulation of immune effector process, positive regulation of cytokine production and neutrophil mediated immunity (Fig. 3B). KEGG pathway analysis mostly included phagosome, complement and coagulation cascades, cell adhesion molecules (CAMs), ECM-receptor interaction and focal adhesion (Fig. 3C). AGE-RAGE signaling pathway in diabetic complications was also found.

miRNA-mRNA network

MicroRNA could inhibit gene expression via base-pairing with mRNA. To construct miRNA-mRNA network, we also analysed features of GSE51674. There were no very heterogeneous samples in sample cluster dendrogram (Fig. 5A). Principal component analysis showed that the two main components respectively contributed 62.71% and 15.68% (Fig. 4B). Next, 16 down regulated DERs and 67 up regulated DERs were found with the criteria of |log2 FC| greater

than 3 and adjusted P-Value less than 0.01. In addition, the volcano map displayed the expression change and down regulated DERs were shown in the hierarchical clustering heatmap (Fig. 4C and 4D). 16 down regulated DERs were used to construd miRNA-mRNA network. 88 miRNA-mRNA pairs were obtained according miRNA target webtools (Fig. 5A) and visualized in (Fig. 5B). Among them, TGFBI, SH2B3 and ZNF652 were up regulated in GSE30529.

PPI network and hub gene identification

To construct PPI network, combined gene list was export to STRING database with highest interacion confidence score 0.9. Cytoscape software was used to present this network containg 190 nodes and 680 edges (Fig. 6A). The size and gradient color of nodes are adjusted by degree,while the thickness and gradient color of edge are adjusted by interaction score. CytoHubba provides 12 methods to search important nodes in networks. Each algorithm respectively computed all genes scores and then 50 to 1 points were assigned based on rank. According to all points, the top 20, KNG1, C3, FN1, SYK, HLA-E, EGF, ITGB2, CXCL1, CXCL8, ITGAV, LYN, VWF, RHOA, HLA-DQA1, ITGAM, SERPING1, P2RY13, ANXA1, P2RY14 and FCER1G were identified in heatmap (Fig. 6B). C3 is particularly noteworthy.

Gene methylation differences

It is generally accepted that genetic and environmental factors are involved in the development of DN. DNA methylation could be the bridge (22). To learn whether the target gene has methylation differences, GSE121820 profiles were downloaded as validation cohort. As a result, 15 hub genes, ANXA1, C3, CXCL1, CXCL8, FCER1G, FN1, HLA-E, ITGAV, ITGB2, KNG1, LYN, P2RY13, P2RY14, RHOA and VWF indeed existed DNA methylation difference between DN group and control group with P-value less than 0.05 (Fig. 6C).

Clinical data validation and drug prediction

To explore potential role of hub genes in DN, Pearson correlation analysis was performed between hub genes and clincal data. Gene expression of SYK, CXCL1, LYN, VWF, ANXA1, C3, HLA-E, RHOA and SERPING1 negatively related with GFR, suggesting pathogenic role of up regulated genes in DN (Fig. 8). Conversely, gene expression of EGF and KNG1 positively related with GFR, suggesting protective role of down-regulated genes in DN (Fig. 7A-8F). SYK, CXCL1, LYN, VWF, ANXA1, C3, HLA-E, RHOA, SERPING1, EGF and KNG1 were defined as target genes.

Given that the effective of existing treatment strategies is not entirely satisfactory, 23 up regulated DEGs (logFC greater than 2.5) and 13 down regulated DEGs (logFC less than 1.5) were used to search potential drugs in Connectivity Map. 10 small molecule compounds were identified as potential therapeutic drugs (Table 1).

Table 1
Changes caused by small molecular compounds in cell lines.

Cell ID Mocule	HEPG2	HA1E	A375	HT29	PC3	HCC515	A549	VCAP	MCF7	summary
VEGF-receptor-2-kinase-inhibitor-IV	NaN	-97.65	NaN	NaN	-98.72	-36.65	NaN	39.31	NaN	-95.76
GPR158	-99.67	-99.54	-87.54	-85.11	-95.37	-96.24	NaN	0.00	-78.44	-99.00
CYP51A1	-96.01	-25.96	-98.71	66.44	-91.52	NaN	-99.55	NaN	0.00	-99.38
PTMS	NaN	0.00	-98.98	-93.33	-78.99	-96.77	-98.10	-35.22	0.00	-97.04
withaferin-a	0.00	0.00	-92.00	-14.15	-97.00	-98.40	0.00	-90.66	-96.15	-95.84
digoxin	NaN	0.00	-98.81	0.00	-76.95	-97.80	0.00	-91.20	-97.67	-96.35
digoxin	0.00	0.00	-97.41	0.00	-96.72	-83.92	-93.40	-98.91	-87.77	-95.77
ouabain	-91.46	-25.32	-96.72	-88.99	-89.54	-53.35	-97.20	0.00	-92.89	-94.13
PHF15	64.27	-81.91	-95.04	0.00	-96.29	NaN	-96.54	NaN	0.00	-95.50
ANKRD10	0.00	-89.60	-22.95	-93.47	0.00	NaN	-83.94	NaN	-88.00	-86.68

Discussion

As one of microvascular complication of diabetes, DN is the main cause of ESRD. Existing treatments are not sufficient to control disease progression. New treatment strategies were needed. High-throughput omics data have been widely used to study mechanisms of disease and predict possible therapeutic targets. Current research aboout DN published on the public platform are mostly generated from a single-cohort study. Therefore, an integrated analysis of the data is needed. We performed differential expression analysis and WGCNA with gene expression profiles. GO analysis of combined gene list mainly included neutrophil activation, regulation of immune effector process, positive regulation of cytokine production and neutrophil mediated immunity. KEGG pathway analysis mostly included phagosome, complement and coagulation cascades, cell adhesion molecules (CAMs), ECM-receptor interaction, focal adhesion and AGE-RAGE signaling pathway in diabetic complications. The results supported that immune and inflammatory response may involve in DN. Cytokine release and extracellular matrix deposition may be subsequent events and continue with development of disease. Next, PPI network was established and the 20 hub genes were identificated from it. We also analysed one miRNA profiles to construct mRNA-miRNA network. The network suggested miR-766-3p/TGFBI, miR-1238-5p/ZNF652 and miR-1237-3p/SH2B3 axis. Because epigenetic modifications are considered an important factor during the long course of DN, 15 genes (ANXA1, C3, CXCL1, CXCL8, FCER1G, FN1, HLA-E, ITGAV, ITGB2, KNG1, LYN, P2RY13, P2RY14, RHOA and VWF) were verified methylation difference. But the degree of change is not attainable from GSE121820. Futhermore, correlation analysis with clinical data demonstrated the disease-promoting effect of SYK, CXCL1, LYN, VWF, ANXA1, C3, HLA-E, RHOA and SERPING1, which were up regulated in DN tubule samples. On the contrary, EGF and KNG1 was proved protective effect in DN, which were down regulated in DN tubule samples.

So far, there have been some reports about target genes and DN. Spleen tyrosine kinase (SYK) was reported to mediate high glucose induced TGF-β1 increasement and IL-1β secretion (23, 24). In two animal experiments, C-X-C motif chemokine ligand 1 (CXCL1) stimulated in DN pathogenic environment may served as proinflmmatory mediator (25, 26). In addition, VWF was reported to involve in intrarenal thrombosis leading to deterioration of renal function (27). Purvis et al. observed higher circulating plasma levels of ANXA1 in T1D and T2D patients, whereas exogenous supplement of ANXA1 improves insulin resistance and keeps off the progression of subsequent microvascular complications in mice (28, 29). Previous studies have demonstrated that statins prevent from DN by reducing Ras homolog family member A (RhoA) protein activation (30–33). Another study reported that activation of RhoA/ROCK may regulate NF-κB signaling pathway (34). In addition, sinomenine, kaempferol, catalpol and rutin have been shown a protective effects through RhoA/ROCK signaling pathway (35–38). EGF was considered as urine biomakers in two researches (39, 40). Recently, a newest report about cytosine methylation differences in kidney tubule samples supported this viewpoint (41). Besides, one large-scale linkage study revealed polymorphisms in kininogen 1 (KNG1) associated with DN in European populations (42).

Results is consistent with knowledge that complement system participates in DN. The development of diabetes intimately linked to low-grade inflammation (43). High levels of inflammatory markers such as c-reactive protein and adiponectin proved this viewpoint (44, 45). Inflammation might promote the occurrence and development of diabetic complications such as DN. But it is still poorly understood about underlying mechanisms of initiation of low-grade inflammation. More and more research evidence proved innate immune system are closely involved in diabetes (46). Simultaneously, roles for pattern recognition receptors (PRRs) have been discussed in related with DN (47, 48). Complement system is not only involved in innate immune defence but also considered as an important proinflammaton factor. Several studies pointed out that the complement system is involved in the pathogenesis of DN and might be a therapeutic target (49–51). Significant difference of complement system components level in both plasma and urine were found between patients sufferd from DN and kidney exempted patients. In addition, Li et al highlighted the relatively more important impact of C3a, C5a and sC5b-9 in the development of DN (52). Sun et al demonstrated more severe kidney damage in renal histopathology assessment was associated with deposition of C1q and C3c (53). Futhermore, a large-scale cohort study substantiated high plasms levels of C3 are more prone to kidney damage in individuals from the general population (54). Another research indicted that serum levels of C3 may help to differentiate non-diabetic renal disease from DN in patients with T2DM (55). Blockade of C3a and C5a receptors in the T1DM model indicated a potential protective effect on renal fibrosis by improving endothelial to myofibroblast transition through the Wnt/β-catenin signalling pathway (56). Similarly, blockade of C3a receptors in rats with T2DM improved renal morphology and function by inhibiting cytokine release and TGFβ/Smad3 signalling (57). However, the best approach targeting the complement system to prevent the development of DN still needs to be explored. Therefore 10 potential small molecule compounds have been identified by Connectivity Map database in our study.

Conclusions

In a word, our study has important significance in understanding the underlying mechanisms and is helpful to develop new treatment strategies. However, further molecular biological experiments are needed to verify the association between identified genes and DN.

Abbreviations

WGCNA: weighted correlation network analysis; GO: gene ontology; PPI: protein-protein interaction; DN: diabetic nephropathy; ESRD: end-stage renal disease; DEG: differentially expressed gene; FC: fold change; DER: differentially expressed miRNA; TOM: Topological overlap matrix; KEGG: Kyoto Encyclopedia of Genes and Genomes; Deg: Degree method; MNC: Maximum Neighborhood Component; DMNC: Density of Maximum Neighborhood Component; MCC: Maximal Clique Centrality; EPC: Edge Percolated Component; CAMs: cell adhesion molecules; SYK: Spleen tyrosine kinase; CXCL1: C-X-C motif chemokine ligand 1; RhoA: Ras homolog family member A; KNG1: kininogen 1; PRRs: pattern recognition receptors;

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

GSE30529, GSE51674 and GSE121820 are available in GEO database (https://www.ncbi.nlm.nih.gov/geo/). The R script data used to support the findings of this study are included within the supplementary information file.

Competing interests

The authors declare that they have no competing interest.

Funding

This work was supported by the National Natural Science Foundation of China (grant no 81770743) and Natural Science Foundation of HuNan Provience of China (grant no 2018JJ2654).

Authors' contributions

ST: takes responsibility for analysis of datas and writing the manuscript. XW, TD and HG: takes responsibility for all aspects of the reliability and freedom from bias of the data presented. XX: takes responsibility for full text evaluation and guidance, final approval of the version to be submitted. All authors read and approved the final manuscript.

Acknowledgements

Not applicable.

References

1. Cho NH, Shaw JE, Karuranga S, Huang Y, da Rocha Fernandes JD, Ohlrogge AW, et al. IDF Diabetes Atlas: Global estimates of diabetes prevalence for 2017 and projections for 2045. Diabetes Res Clin Pract. 2018;138:271-81.
2. Zeni L, Norden AGW, Cancarini G, Unwin RJ. A more tubulocentric view of diabetic kidney disease. J Nephrol. 2017;30(6):701-17.

3. Ruggenenti P, Cravedi P, Remuzzi G. The RAAS in the pathogenesis and treatment of diabetic nephropathy. Nat Rev Nephrol. 2010;6(6):319-30.
4. Petryszak R, Burdett T, Fiorelli B, Fonseca NA, Gonzalez-Porta M, Hastings E, et al. Expression Atlas update--a database of gene and transcript expression from microarray- and sequencing-based functional genomics experiments. Nucleic Acids Res. 2014;42(Database issue):D926-32.
5. Woroniecka KI, Park AS, Mohtat D, Thomas DB, Pullman JM, Susztak K. Transcriptome analysis of human diabetic kidney disease. Diabetes. 2011;60(9):2354-69.
6. 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. 2017;7(1):8576.
7. Ju W, Greene CS, Eichinger F, Nair V, Hodgin JB, Bitzer M, et al. Defining cell-type specificity at the transcriptional level in human disease. Genome research. 2013;23(11):1862-73.
8. Grayson PC, Eddy S, Taroni JN, Lightfoot YL, Mariani L, Parikh H, et al. Metabolic pathways and immunometabolism in rare kidney diseases. Annals of the rheumatic diseases. 2018;77(8):1226-33.
9. Conserva F, Barozzino M, Pesce F, Divella C, Oranger A, Papale M, et al. Urinary miRNA-27b-3p and miRNA-1228-3p correlate with the progression of Kidney Fibrosis in Diabetic Nephropathy. Sci Rep. 2019;9(1):11357.
10. Davis S, Meltzer PS. GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. Bioinformatics. 2007;23(14):1846-7.
11. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;43(7):e47.
12. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008;9:559.
13. Agarwal V, Bell GW, Nam JW, Bartel DP. Predicting effective microRNA target sites in mammalian mRNAs. Elife. 2015;4.
14. Sticht C, De La Torre C, Parveen A, Gretz N. miRWalk: An online resource for prediction of microRNA binding sites. PLoS One. 2018;13(10):e0206239.
15. Kozomara A, Birgaoanu M, Griffiths-Jones S. miRBase: from microRNA sequences to function. Nucleic Acids Res. 2019;47(D1):D155-D62.
16. Chou CH, Shrestha S, Yang CD, Chang NW, Lin YL, Liao KW, et al. miRTarBase update 2018: a resource for experimentally validated microRNA-target interactions. Nucleic Acids Res. 2018;46(D1):D296-D302.
17. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS. 2012;16(5):284-7.
18. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. 2019;47(D1):D607-D13.
19. Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY. cytoHubba: identifying hub objects and sub-networks from complex interactome. BMC Syst Biol. 2014;8 Suppl 4:S11.
20. Schmid H, Boucherot A, Yasuda Y, Henger A, Brunner B, Eichinger F, et al. Modular activation of nuclear factor-kappaB transcriptional programs in human diabetic nephropathy. Diabetes. 2006;55(11):2993-3003.
21. Subramanian A, Narayan R, Corsello SM, Peck DD, Natoli TE, Lu X, et al. A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles. Cell. 2017;171(6):1437-52 e17.

22. Singh S, Sonkar SK, Sonkar GK, Mahdi AA. Diabetic kidney disease: A systematic review on the role of epigenetics as diagnostic and prognostic marker. Diabetes/Metabolism Research and Reviews. 2019;35(5).
23. Yang WS, Chang JW, Han NJ, Lee SK, Park SK. Spleen tyrosine kinase mediates high glucose-induced transforming growth factor-beta1 up-regulation in proximal tubular epithelial cells. Exp Cell Res. 2012;318(15):1867-76.
24. Qiao Y, Tian X, Men L, Li S, Chen Y, Xue M, et al. Spleen tyrosine kinase promotes NLR family pyrin domain containing 3 inflammasomemediated IL1beta secretion via cJun Nterminal kinase activation and cell apoptosis during diabetic nephropathy. Mol Med Rep. 2018;18(2):1995-2008.
25. Niu S, Bian Z, Tremblay A, Luo Y, Kidder K, Mansour A, et al. Broad Infiltration of Macrophages Leads to a Proinflammatory State in Streptozotocin-Induced Hyperglycemic Mice. J Immunol. 2016;197(8):3293-301.
26. Xu J, Zheng S, Kralik PM, Krishnan L, Huang H, Hoying JB, et al. Diabetes Induced Changes in Podocyte Morphology and Gene Expression Evaluated Using GFP Transgenic Podocytes. Int J Biol Sci. 2016;12(2):210-8.
27. Dhanesha N, Doddapattar P, Chorawala MR, Nayak MK, Kokame K, Staber JM, et al. ADAMTS13 Retards Progression of Diabetic Nephropathy by Inhibiting Intrarenal Thrombosis in Mice. Arteriosclerosis, thrombosis, and vascular biology. 2017;37(7):1332-8.
28. Purvis GSD, Chiazza F, Chen J, Azevedo-Loiola R, Martin L, Kusters DHM, et al. Annexin A1 attenuates microvascular complications through restoration of Akt signalling in a murine model of type 1 diabetes. Diabetologia. 2018;61(2):482-95.
29. Purvis GSD, Collino M, Loiola RA, Baragetti A, Chiazza F, Brovelli M, et al. Identification of AnnexinA1 as an Endogenous Regulator of RhoA, and Its Role in the Pathophysiology and Experimental Therapy of Type-2 Diabetes. Front Immunol. 2019;10:571.
30. Danesh FR, Sadeghi MM, Amro N, Philips C, Zeng L, Lin S, et al. 3-Hydroxy-3-methylglutaryl CoA reductase inhibitors prevent high glucose-induced proliferation of mesangial cells via modulation of Rho GTPase/ p21 signaling pathway: Implications for diabetic nephropathy. Proc Natl Acad Sci U S A. 2002;99(12):8301-5.
31. Zeng L, Xu H, Chew TL, Eng E, Sadeghi MM, Adler S, et al. HMG CoA reductase inhibition modulates VEGF-induced endothelial cell hyperpermeability by preventing RhoA activation and myosin regulatory light chain phosphorylation. Faseb j. 2005;19(13):1845-7.
32. Kolavennu V, Zeng L, Peng H, Wang Y, Danesh FR. Targeting of RhoA/ROCK signaling ameliorates progression of diabetic nephropathy independent of glucose control. Diabetes. 2008;57(3):714-23.
33. Peng F, Wu D, Gao B, Ingram AJ, Zhang B, Chorneyko K, et al. RhoA/Rho-kinase contribute to the pathogenesis of diabetic renal disease. Diabetes. 2008;57(6):1683-92.
34. Xie X, Peng J, Chang X, Huang K, Huang J, Wang S, et al. Activation of RhoA/ROCK regulates NF-kappaB signaling pathway in experimental diabetic nephropathy. Mol Cell Endocrinol. 2013;369(1-2):86-97.
35. Yin Q, Xia Y, Wang G. Sinomenine alleviates high glucose-induced renal glomerular endothelial hyperpermeability by inhibiting the activation of RhoA/ROCK signaling pathway. Biochem Biophys Res Commun. 2016;477(4):881-6.
36. Wang X, Zhao X, Feng T, Jin G, Li Z. Rutin Prevents High Glucose-Induced Renal Glomerular Endothelial Hyperpermeability by Inhibiting the ROS/Rhoa/ROCK Signaling Pathway. Planta Med. 2016;82(14):1252-7.
37. Sharma D, Gondaliya P, Tiwari V, Kalia K. Kaempferol attenuates diabetic nephropathy by inhibiting RhoA/Rho-kinase mediated inflammatory signalling. Biomed Pharmacother. 2019;109:1610-9.
38. Chen Y, Liu Q, Shan Z, Mi W, Zhao Y, Li M, et al. Catalpol Ameliorates Podocyte Injury by Stabilizing Cytoskeleton and Enhancing Autophagy in Diabetic Nephropathy. Front Pharmacol. 2019;10:1477.

39. Satirapoj B, Dispan R, Radinahamed P, Kitiyakara C. Urinary epidermal growth factor, monocyte chemoattractant protein-1 or their ratio as predictors for rapid loss of renal function in type 2 diabetic patients with diabetic kidney disease. BMC Nephrol. 2018;19(1):246.
40. Nowak N, Skupien J, Smiles AM, Yamanouchi M, Niewczas MA, Galecki AT, et al. Markers of early progressive renal decline in type 2 diabetes suggest different implications for etiological studies and prognostic tests development. Kidney Int. 2018;93(5):1198-206.
41. Gluck C, Qiu C, Han SY, Palmer M, Park J, Ko Y-A, et al. Kidney cytosine methylation changes improve renal function decline estimation in patients with diabetic kidney disease. Nature Communications. 2019;10(1).
42. Vionnet N, Tregouet D, Kazeem G, Gut I, Groop PH, Tarnow L, et al. Analysis of 14 candidate genes for diabetic nephropathy on chromosome 3q in European populations: strongest evidence for association with a variant in the promoter region of the adiponectin gene. Diabetes. 2006;55(11):3166-74.
43. Lontchi-Yimagou E, Sobngwi E, Matsha TE, Kengne AP. Diabetes mellitus and inflammation. Curr Diab Rep. 2013;13(3):435-44.
44. Mazidi M, Toth PP, Banach M. C-reactive Protein Is Associated With Prevalence of the Metabolic Syndrome, Hypertension, and Diabetes Mellitus in US Adults. Angiology. 2018;69(5):438-42.
45. Bruun JM, Lihn AS, Verdich C, Pedersen SB, Toubro S, Astrup A, et al. Regulation of adiponectin by adipose tissue-derived cytokines: in vivo and in vitro investigations in humans. American journal of physiology Endocrinology and metabolism. 2003;285(3):E527-33.
46. Wada J, Makino H. Innate immunity in diabetes and diabetic nephropathy. Nat Rev Nephrol. 2016;12(1):13-26.
47. Mudaliar H, Pollock C, Panchapakesan U. Role of Toll-like receptors in diabetic nephropathy. Clin Sci (Lond). 2014;126(10):685-94.
48. Du P, Fan B, Han H, Zhen J, Shang J, Wang X, et al. NOD2 promotes renal injury by exacerbating inflammation and podocyte insulin resistance in diabetic nephropathy. Kidney Int. 2013;84(2):265-76.
49. Fortpied J, Vertommen D, Van Schaftingen E. Binding of mannose-binding lectin to fructosamines: a potential link between hyperglycaemia and complement activation in diabetes. Diabetes Metab Res Rev. 2010;26(4):254-60.
50. Acosta J, Hettinga J, Fluckiger R, Krumrei N, Goldfine A, Angarita L, et al. Molecular basis for a link between complement and the vascular complications of diabetes. Proc Natl Acad Sci U S A. 2000;97(10):5450-5.
51. Flyvbjerg A. The role of the complement system in diabetic nephropathy. Nat Rev Nephrol. 2017;13(5):311-8.
52. Li XQ, Chang DY, Chen M, Zhao MH. Complement activation in patients with diabetic nephropathy. Diabetes Metab. 2019;45(3):248-53.
53. Sun ZJ, Li XQ, Chang DY, Wang SX, Liu G, Chen M, et al. Complement deposition on renal histopathology of patients with diabetic nephropathy. Diabetes Metab. 2019;45(4):363-8.
54. Rasmussen KL, Nordestgaard BG, Nielsen SF. Complement C3 and Risk of Diabetic Microvascular Disease: A Cohort Study of 95202 Individuals from the General Population. Clin Chem. 2018;64(7):1113-24.
55. Zhang J, Wang Y, Zhang R, Li H, Han Q, Wu Y, et al. Serum levels of immunoglobulin G and complement 3 differentiate non-diabetic renal disease from diabetic nephropathy in patients with type 2 diabetes mellitus. Acta Diabetol. 2019;56(8):873-81.
56. Li L, Chen L, Zang J, Tang X, Liu Y, Zhang J, et al. C3a and C5a receptor antagonists ameliorate endothelial-myofibroblast transition via the Wnt/beta-catenin signaling pathway in diabetic kidney disease. Metabolism. 2015;64(5):597-610.

57. Li L, Yin Q, Tang X, Bai L, Zhang J, Gou S, et al. C3a receptor antagonist ameliorates inflammatory and fibrotic signals in type 2 diabetic nephropathy by suppressing the activation of TGF-beta/smad3 and IKBalpha pathway. PLoS One. 2014;9(11):e113639.

Figures

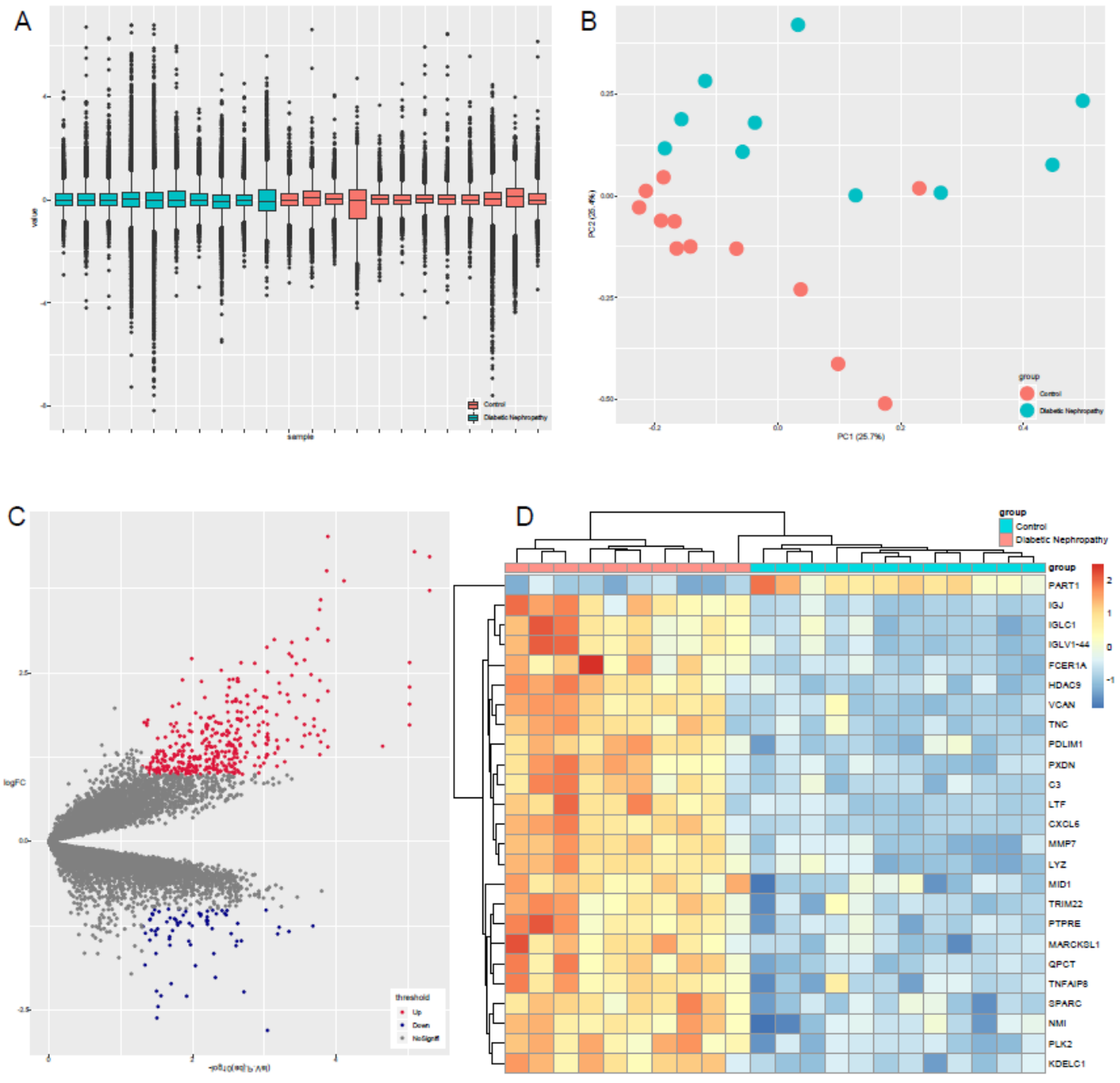

Figure 1

a) Boxplot of GSE30529. (b) Principal component analysis of GSE30529. The two main components respectively contributed 25.7% and 25.4%. (c) Volcano map of DEGs. Genes with log2 FC greater than 1 and adjusted P-Value less

than 0.05 was thought as up regulated gene. Genes with log2 FC less than -1 and adjusted P-Value less than 0.05 was thought as down regulated gene. (d) Heatmap of top 25 DEGs.

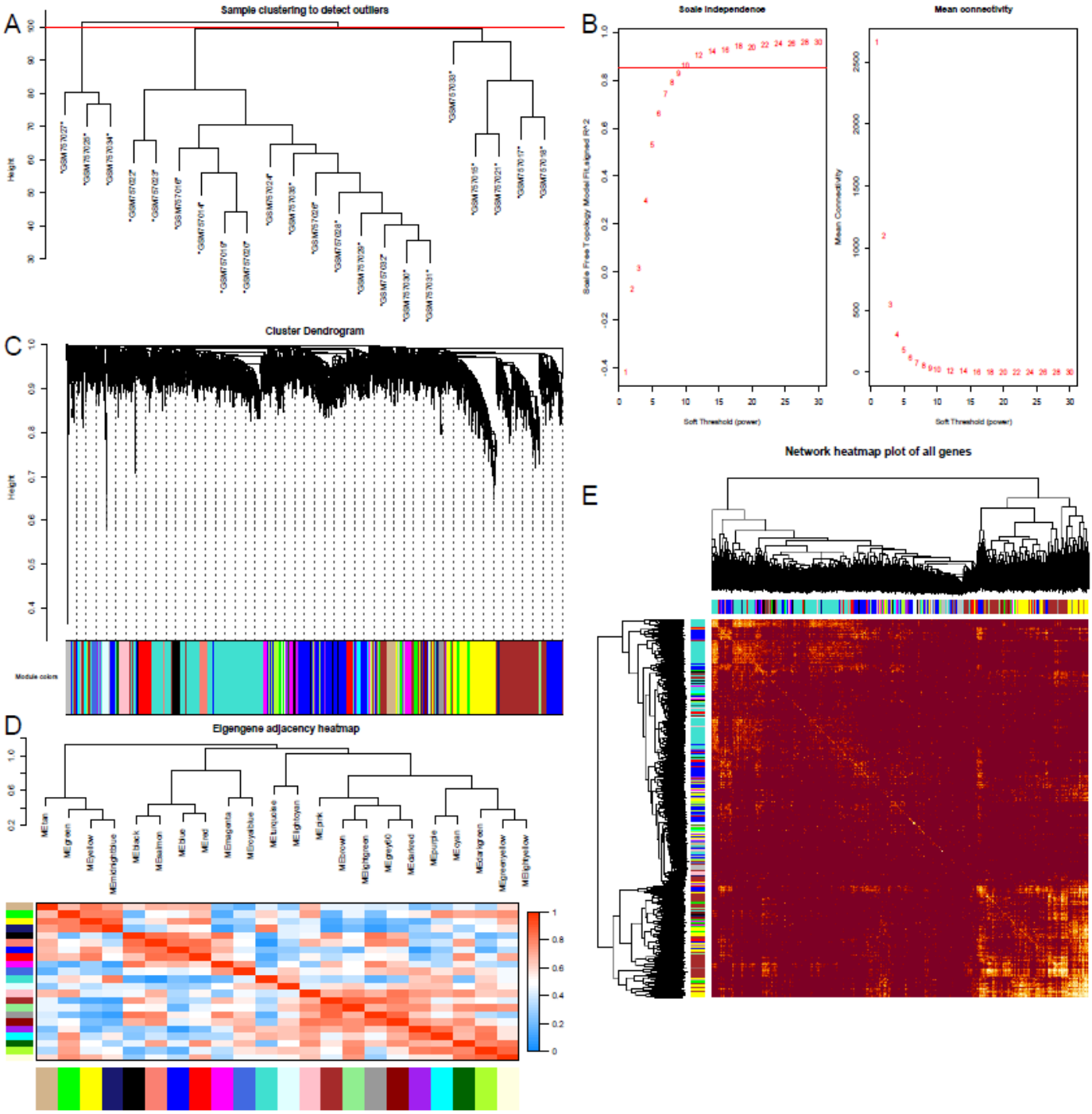

Figure 2

(a) Sample Clustering of GSE30529. (b) Analysis of soft-thresholding powers and mean connectivity. 10 is chosen from among them to construct a scale free network. (c) Dendrogram of the gene modules. The branches represent different gene modules and each leaf represents a gene in cluster dendrogram. (d) Clustering and heatmap of 22 gene modules. (e) Weighted correlations of all genes and clustering.

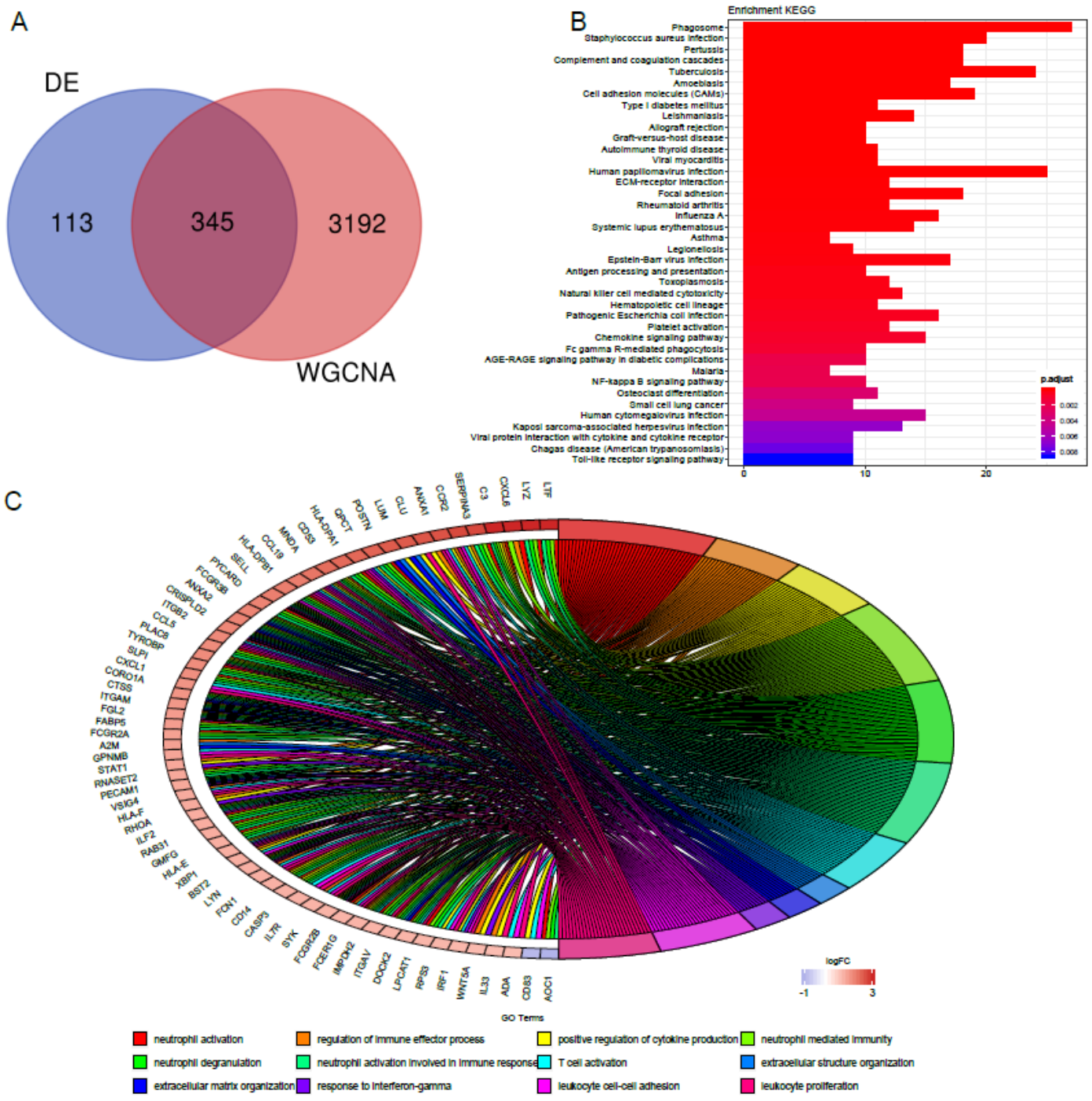

Figure 3

(a) Venn plot of DEG list and highly related gene list. 345 genes were obtained. (b and c) GO annotation and KEGG pathway enrichment analysis. GO analysis mainly included neutrophil activation, regulation of immune effector process, positive regulation of cytokine production and neutrophil mediated immunity. KEGG pathway analysis mostly included phagosome, complement and coagulation cascades, cell adhesion molecules and ECM-receptor interaction and focal adhesion.

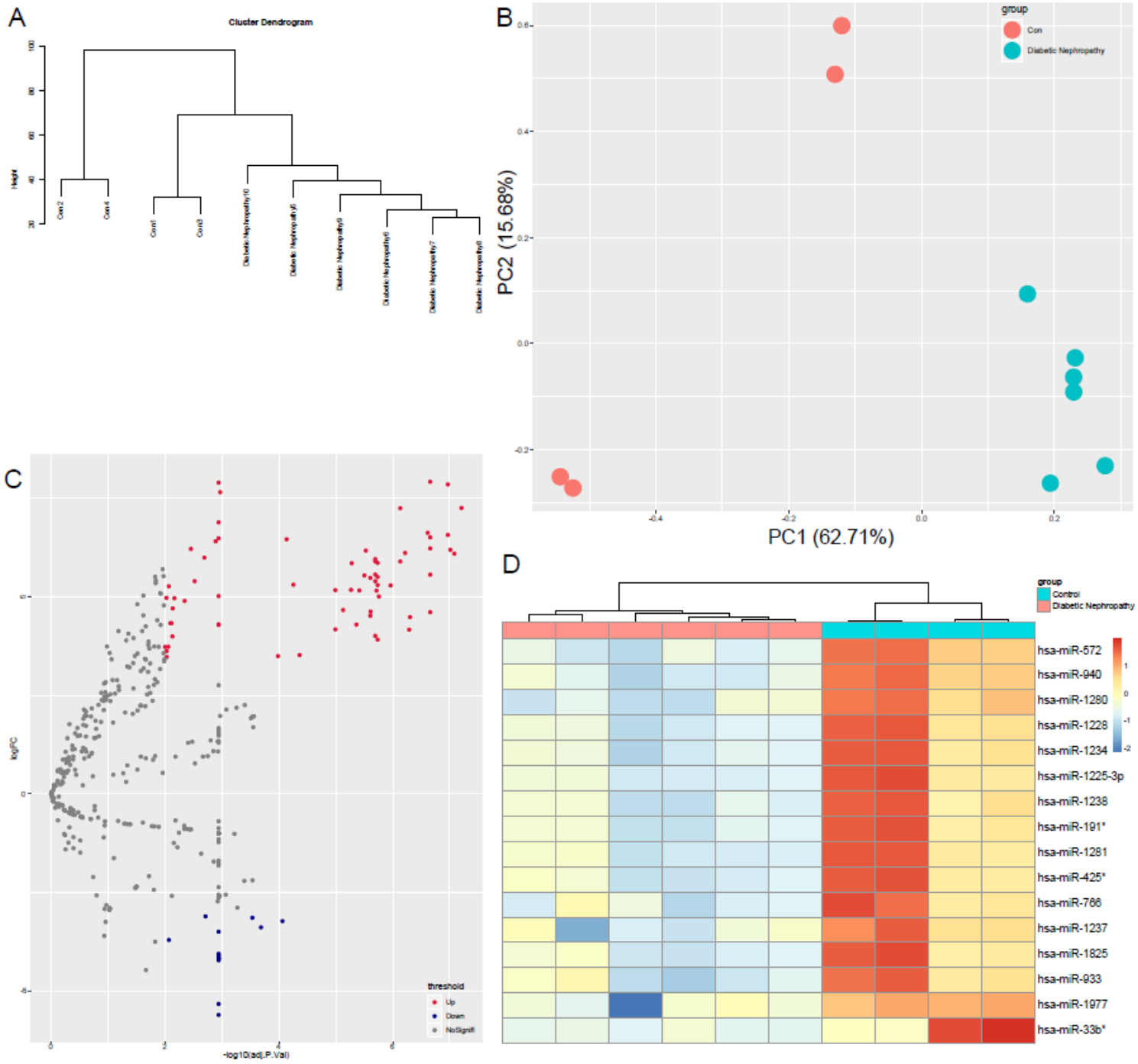

Figure 4

(a) cluster dendrogram of GSE51674. (b) Principal component analysis of GSE30529. The two main components respectively contributed 62.71% and 15.68%. (c) Volcano map of DERs. MicroRNAs with log2 FC greater than 3 and adjusted P-Value less than 0.01 was thought as up regulated DERs. MicroRNAs with log2 FC less than -3 and adjusted P-Value less than 0.01 was thought as down regulated DERs. (d) Heatmap of down regulated DERs.

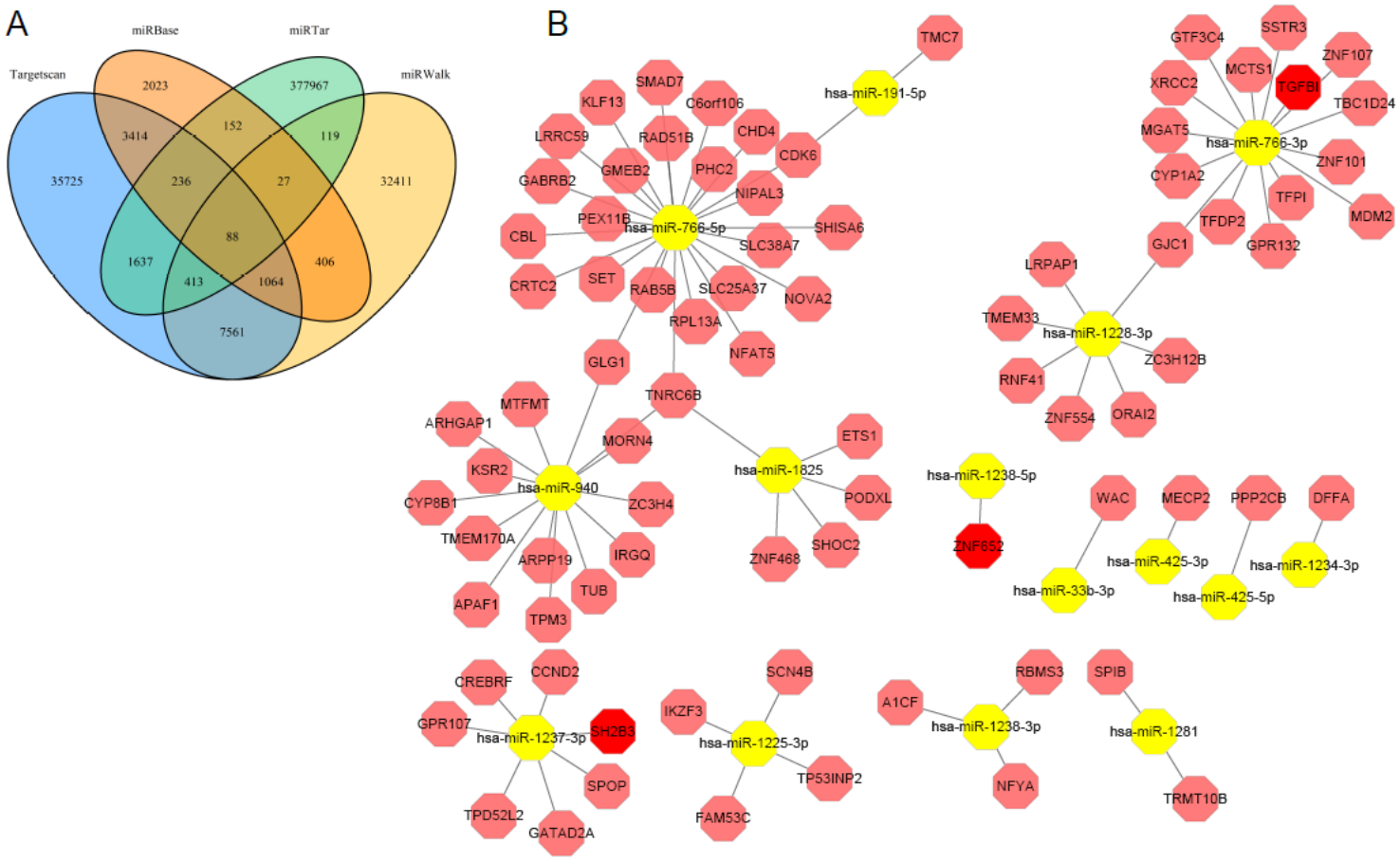

Figure 5

(a) Venn plot of four prediction results. (b). miRNA-mRNA network. Among them, TGFBI, SH2B3 and ZNF652 in red were up regulated in GSE30529.

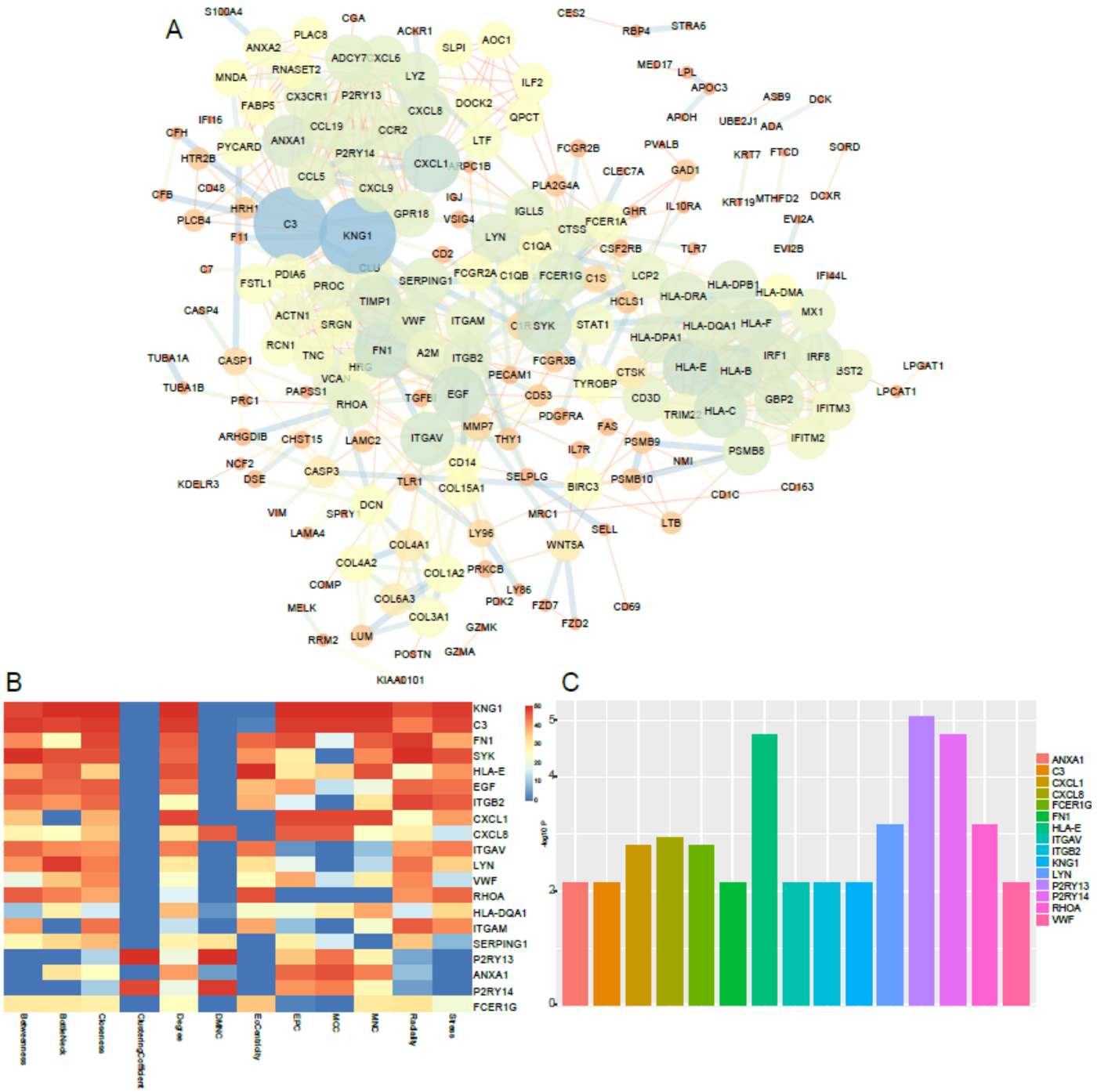

Figure 6

(a) PPI network of combined genes.There are 190 nodes and 680 edges. The size and gradient color of nodes are adjusted by degree. Thickness and gradient color of edge are adjusted by interaction score. (b) Heatmap of CytoHubba analysis score. (c) DNA methylation difference of hub genes in DN.

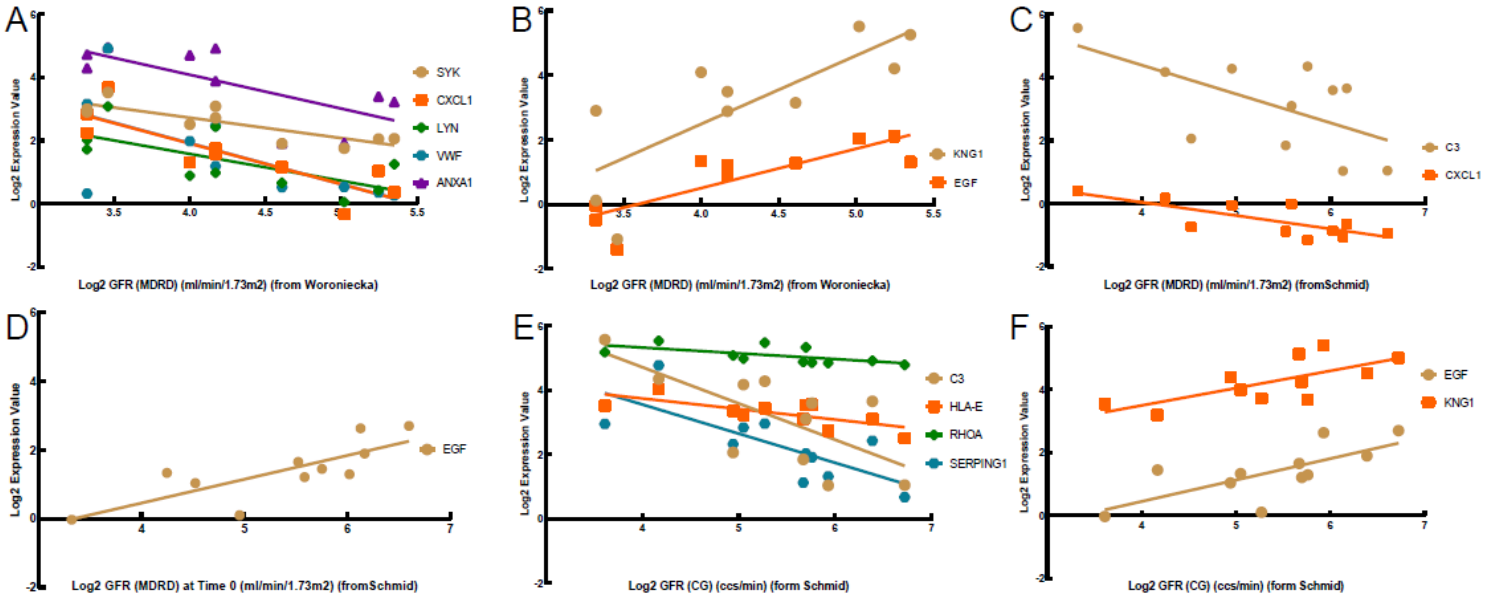

Figure 7

Pearson correlation analyses of GFR and target genes. (a) Gene expression of SYK (p=0.0022, r=-0.8437), CXCL1 (p=0.0016, r=-0.8554), LYN (p=0.0269, r=-0.6911), VWF (p=0.0452, r=-0.6423) and ANXA1 (p=0.0211, r=-0.7111) negatively related with GFR. (b) Gene expression of EGF (p=0.0027, r=0.8349) and KNG1 (p=0.0073, r=0.7838) positively related with GFR. (c) Gene expression of C3 (p=0.0459, r=-0.6109) and CXCL1 (p=0.0061, r=-0.7645) negatively related with GFR. (d) Gene expression of EGF (p=0.0037, r=0.7919) positively related with GFR. (e) Gene expression of C3 (p=0.0171, r=-0.6970), HLA-E (p=0.0132, r=-0.7161), RHOA (p=0.0439, r=-0.6154) and SERPING1 (p=0.0091, r=-0.7409) negatively related with GFR. (f) EGF (p=0.0121, r=0.7221) and KNG1 (p=0.0153, r=0.7053) positvely related with GFR.