Prediction of Molecular Signature, Potential Biomarkers, and Molecular Pathways Associated with Membranous Nephropathy Based on Protein-Protein Interactions

Amir Taherkhani¹, Shiva Kalantari²* and Mohsen Nafar²

¹Department of Basic Science, Faculty of Paramedical Sciences, and ²Chronic Kidney Disease Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran

ABSTRACT

Background: Membranous nephropathy (MN) is one of the causes of nephrotic syndrome in adults that lead to end-stage renal disease with an unknown molecular signature. The current diagnosis is based on renal biopsy, which is an invasive method and has several complications and challenges. Thus, identification of the novel biomarker candidates, as well as impaired pathways, will be helpful for non-invasive molecular-based diagnosis. Objectives: We aimed to study the molecular signature of MN and facilitate the systematic discovery of diagnostic candidate biomarkers, molecular pathway, and potential therapeutic targets using bioinformatics predictions. Methods: The protein-protein interaction (PPI) network of an integrated list of downloaded microarray data, differential proteins from a published proteomic study, and a list of retrieved scientific literature mining was constructed and analyzed in terms of functional modules, enriched biological pathways, hub genes, master regulator, and target genes. Results: These network analyses revealed several functional modules and hub genes including Vitamin D3 receptor, retinoic acid receptor RXR-alpha, interleukin 8, and SH3GL2. TEAD4 and FOXA1 were identified as the regulatory master molecules. LRP1 and ITGA3 were identified as the important target genes. Extracellular matrix organization, cell surface receptor signaling pathway, and defense and inflammatory response were found to be impaired in MN using functional analyses. A specific subnetwork for MN was suggested using PPI approach. Discussion: Omics data integration and systems biology analysis on the level of interaction networks provide a powerful approach for identification of pathway-specific biomarkers for MN. (REV INVES CLIN. 2018;70:184-91)


Corresponding author:
*Shiva Kalantari
Chronic Kidney Disease Research Center
Shahid Beheshti University of Medical Sciences
Pasdaran Street, Boostan 9th, Number 103
Tehran, Iran. Postcode: 1666663111
E-mail: shiva.kalantari@sbmu.ac.ir

Received for publication: 03-04-2108
Accepted for publication: 12-06-2018
doi: 10.24875/RIC.18002551
INTRODUCTION

Membranous nephropathy (MN), an immunologically mediated disease, is one of the causes of nephrotic syndrome in adults over 60 years old\(^1\)\(^2\). MN could be either primary (75% of the cases) or secondary. The etiology of the primary cases is idiopathic, while the secondary form is associated with various conditions such as hepatitis infection, malignancy, systemic lupus erythematosus and also drug intoxication\(^3\)\(^4\). It is characterized by the glomerular basement membrane thickening due to subepithelial deposition of immune complexes (immunoglobulin G and complement factors), which causes podocyte damage and usually nephrotic syndrome\(^3\)\(^4\). Decline in renal function and development of end-stage renal disease occur in approximately 40% of the patients with idiopathic MN\(^4\). Therefore, early diagnosis using a molecular signature of the disease or molecular biomarkers will improve kidney survival and reduces the therapeutic toxicity. The current diagnosis is based on renal biopsy, which is an invasive method and has several complications and challenges. Thus, alternative non-invasive molecular-based diagnosis through the identification (discovery) of novel biomarker candidates is highly needed.

Several suggested biomarkers for idiopathic type of MN (IMN) such as anti-phospholipase A2 receptor (PLA2R) antibody and anti-neutral endopeptidase antibody did not hold reliable sensitivity and specificity\(^5\)\(^6\). Although PLA2R may be a valuable marker for disease monitoring and response to therapy, it may not be good enough for diagnosis because circulating antibodies against PLA2R are found in up to 70-80% of patients with idiopathic MN, while the remaining patients are negative for this antibody\(^7\). Besides anti-PLA2R, the other antibodies against autoantigens (e.g., thrombospondin type-1 domain-containing 7A protein [THSD7A], superoxide dismutase [SOD2], and aldose reductase) that are suggested as biomarkers for primary MN have been reviewed elsewhere\(^8\). Several of these suggested candidates need more validation.

Despite these progress made for improvement of diagnosis and evaluation of the patient’s status, the autoimmune basis of MN is not clearly understood. Thus, further research into pathogenic mechanisms and other disease markers are still needed for improving patient care.

Systems biology has the potential to extract the meaning from superabundant and inexpressive data derived from omics technologies\(^8\). It is an encouraging approach to reduce restrictions in the field of diagnosis and makes progress in the understanding of the mechanisms involved in the pathogenesis of diseases.

In the present study, we applied a systems biology approach to the analysis of protein-protein interaction (PPI) network for prediction of potential candidate biomarkers of MN, understanding of the impaired biological pathways, and proposing the master regulatory molecules as therapeutic targets. The identified genes are promising candidates for the future targeted \textit{in vivo} experiments for validation. To the best of our knowledge, this work is the first comprehensive \textit{in silico} study of MN using bioinformatics tools.

METHODS

Data collection

We used an integrative list of genes involved in MN from three sources to increase coverage: (i) a gene expression dataset, (ii) a differential protein list from a proteomic study, and (iii) literature mining (Table S1). Details of these sources and extracted differentially expressed genes (DEGs) may be found in Supporting Information file 1.

Construction of PPI network

Using Cytoscape software version 3.2.1\(^10\), a PPI network was constructed for the integrated dataset from DEGs in the microarray, proteomic, and literature mining sources. The network constructed against each database was merged (nodes = 18,725 and edges = 47,233). A subset of this network with human nodes only was applied for further analysis.

Network clustering

Clustering with overlap neighborhood expansion (ClusterONE)\(^11\) was used within the unweighted constructed network to discover densely connected and possibly overlapping regions. In PPI networks, these dense regions are usually considered as protein complexes or fractions of them. ClusterONE expands the
small seeds to obtain these dense regions using a quality function. Groups with the degree > 8 and the minimum density > 0.05 were identified as modules. Clusters with $p < 0.001$ were considered statistically significant. These thresholds were selected based on empirical investigations for obtaining the best relevant results with the least missing information. Subsequently, a new subnetwork was constructed by merging the significant clusters and presented as the functional PPI subnetwork for MN. Further, analyses were performed on this new subnetwork for identification of the hub nodes. The hub nodes were identified in the new subnetwork based on these criteria: (i) betweenness centrality more than average, (ii) closeness centrality more than average, and (iii) the degree of more than 2-fold on average.

Functional enrichment and pathway analysis

To understand the biological relevance of the significant cluster genes, we performed functional enrichment analysis, including Gene Ontology (GO) of biological process (BP) and pathway enrichment analyses against Kyoto Encyclopedia of Genes and Genomes (KEGG) using ClueGO plug-in of Cytoscape\textsuperscript{12}.

Master regulator and effector analysis

We used iRegulon to characterize motifs and regulatory nodes underlying the identified hubs and identify downstream transcription factor (TF) effectors of top suggested hub nodes\textsuperscript{13}. Motif detection in this method is based on a ranking-and-recovery strategy whereby every gene in the human genome is scored by a motif discovery step and TF enrichment is calculated for each set of genes\textsuperscript{14}. A normalized enrichment score (NES) is a parameter for judging the significance of motifs and is computed as the area under the curve (AUC) value of the motif/track minus the mean of all AUCs for all motifs (or tracks) and divided by the standard deviation of all AUCs.

RESULTS

Network analysis

The data from three sources (microarray, proteomic study, and literature mining) were integrated. A total of 224 genes were used. The integrated dataset was mapped into PPIs, thus a PPI network was constructed with 7628 nodes and 28,920 edges. 52 clusters were significant ($p < 0.001$) and applied for construction of a new subnetwork containing 1971 nodes. Top two clusters are shown in figure S1. 37 genes were identified as hubs in the new subnetwork and are listed in table S2.

Functional annotation of the clusters

Overall, 174 GO terms were significantly enriched after removing the redundant terms (Table S3). The top four GO terms (BP) were as follows: extracellular matrix organization ($P = 310 \times 10^{-57}$), cell surface receptor signaling pathway ($P = 570 \times 10^{-39}$), defense response ($P = 1.5 \times 10^{-36}$), and inflammatory response ($P = 1.5 \times 10^{-36}$). In total, 30 pathways enriched, of which top three dysregulated pathways in MN were pathways associated with extracellular matrix (ECM)-receptor interaction ($P = 5.4 \times 10^{-25}$), chemokine signaling pathway ($P = 1.2 \times 10^{-15}$), and complement and coagulation cascades ($P = 5.4 \times 10^{-12}$) (Fig. S2). Details of the enriched KEGG pathways are shown in table S4.

Master regulator of important hubs and effector analysis

To reveal the molecular mechanism involved in the regulation of the top four hub genes (Vitamin D3 receptor [VDR], retinoic acid receptor RXR-alpha [RXRA], interleukin 8 [IL-8], and SH3G2) expression, we identified master regulators of the coexpressed genes using iRegulon\textsuperscript{13}. 62 transcriptional regulators for the three genes (i.e., VDR, RXRA, and IL-8, since no regulator was identified for SH3G2) were identified with an NES > 3. The regulator list is tabulated in table S5. The six common regulators which control these three hubs are shown in figure 1 (white triangles). TEAD4 and FOXA1 had the highest NES (> 6) and are suggested as the most important master regulators in the development of MN, which are analyzed in the discussion section. In addition, downstream effectors of top TF hub nodes (VDR and RXRA) were identified, of which the top 20 target genes are shown in figure 2. Nine targets of these genes are common for both TFs and proposed as the important genes in developing MN.
DISCUSSION

Systems biology approaches and bioinformatics tools have recently become popular in the context of biomarker discovery in the clinical setting. The advantage of this approach is to predict new target molecules (e.g., as biomarkers or drug targets) based on logical-mathematical relationships between the targets and available databases with high probability. These in silico predicted target molecules could then be experimentally examined, thus reducing the error rate of results. In this study, we aimed to predict new

Figure 1. Interaction between transcription factors (regulators) and top three hub genes (Vitamin D3 receptor, retinoic acid receptor RXR-alpha, and interleukin 8). The white triangles represent regulators that are common for the three hubs; the open ellipses represent the other regulator for each of the hubs, the dark ellipses represent hub genes, and the directed edges represent the direction of interaction from regulator to target genes.
biomarkers and pathways associated with MN using a protein network-based approach. These new biomarkers, hub molecules, can be key mediators in the disease course with a computational rationale behind them. Therefore, the suggested novel hubs, after validation, have the potential to be used as an alternative for current invasive diagnostic methods and to elucidate the pathogenesis of the disease.

A total of 224 gene signatures for MN were selected using the integration of three datasets from a microarray study, a comprehensive proteomic study, and an extensive literature mining. The functional clusters and protein complexes were identified in the integrated list of genes. We found that 52 functional clusters encompassing 1971 genes, which were mainly involved in extracellular matrix organization, cell surface receptor signaling pathway, defense response, and inflammatory response, contributed to MN development. Enrichment of a fairly high number of significant clusters under stringent statistical tests indicates the quality of our original dataset. According to table S3, most GO terms refer to impairment of the immune system, which corresponds to the characteristic of primary MN as an autoimmune disease.

We found several significant pathways (ECM-receptor interaction, focal adhesion, and chemokine signaling pathway) and 37 hub genes by PPI network. Top four hub genes including VDR, IL-8 (CXCL8/IL-8), RXRA, and endophilin-A1 (CNSA2, SH3D2A, and SH3G2) were identified by network clustering analyses and are focused for further discussion and analysis of their regulators and target genes. Comparing our findings with the study by Hauser et al. in an experimental model of MN, most of the pathways were found in common, including ECM-receptor interaction, regulation of actin cytoskeleton, MAPK signaling pathway, regulation of cell communication, a few pathways related to cell junctions, a pathway related to neuronal development, transforming growth factor-beta receptor signaling pathway, and toll-like receptor signaling pathway. This agreement between both studies may reinforce the significance of these pathways in the pathophysiology of MN. Hauser et al. suggested that extracellular matrix proteins contribute to the pathophysiology of MN through their role in podocyte loss and influence on monocyte migration into renal compartments, exacerbating the immune response after the initial injury.

VDR is a nuclear hormone receptor for Vitamin D3 and belongs to the family of transcriptional regulatory factors. Effectors of this nuclear hormone receptor regulate a variety of metabolic pathways such as those involved in the immune response. It is believed that Vitamin D deficiency is associated with proteinuria and agonists of Vitamin D could have a therapeutic application for autoimmune disorders and transplantation. In our DEGs extracted from GEO...

Figure 2. Downstream effectors of top transcription factor hub nodes. Vitamin D3 receptor and retinoic acid receptor RXR-alpha are transcription factors and have various target genes (ellipse). Nine target genes (rectangles) are regulated by both transcription factors.
dataset, VDR gene was downregulated in the renal biopsy of MN patients in comparison with normal kidney tissue, which is in agreement with the defective role of this multifunctional molecule in renal disease reported in the literature\textsuperscript{20}. Interestingly, two of our top suggested hubs (VDR and RXRA) were known nuclear hormone receptors that contribute to autoimmune kidney disease and chronic glomerulonephritis\textsuperscript{21-23}. It has been postulated that formation of an in situ antibody-mediated immune complex in the glomerular epithelium plays a crucial role in the initiation of renal injury in MN, and hence, MN is considered an autoimmune kidney disease\textsuperscript{24}. Therefore, impaired functions of these two TFs which could regulate immunity may be the critical cause of MN.

Retinoid X receptors (RXRs) bind to retinoic acid through dimerization with the RAR family (e.g., RAR-\alpha, beta [RAR-\beta], and gamma [RAR-\gamma])\textsuperscript{25}. Therefore, it can be postulated that the function of RXRA is in line with the function of RAR-\alpha. It is known that naive CD4+ T cells basally express the genes encoding RAR\alpha (RARa) and RAR\alpha activates these cells through an unclear mechanism\textsuperscript{26,27}. Based on this evidence, it is suggested here that RXRA contributes to developing the IMN through activation of CD4+ T cells (the mature T-helper cell). This finding corresponds with the experimental evidence of increased levels of CD4+ in IMN patients, which resulted in upregulation of IL-10 and IL-13 (cytokine products of type 2 T-helper cells) in these patients compared with normal individuals\textsuperscript{28,29}. Kuroki et al. suggested the aggravating effect of Th2 cytokines on B cells from several IMN patients by increasing the production of IgG4, a predominant subtype of IgG, in IMN\textsuperscript{28}.

IL-8/CXCL8, one of the important hubs in our study, is a member of the CXC chemokine subfamily which predominantly attracts neutrophils to the site of acute and chronic inflammation\textsuperscript{30-33}. However, Lionaki et al. did not detect significant changes in the serum levels of IL-8 in patients with MN, although Niemir et al. revealed upregulation of IL-8 mRNA in the renal biopsy of several glomerular diseases including MN\textsuperscript{34,35}.

The fourth important hub in the present study was endophilin-A1, a member of regulatory components of clathrin-coated vesicle formation\textsuperscript{36}. Endophilin-A1 is required for endocytosis, which controls internalization of micronutrients and turnover of membrane components\textsuperscript{37}. The relationship between endophilin-A1 and renal disease comes from the role of this protein in uptake of several receptor tyrosine kinases including EGFR, VEGFR, PDGFR, and IGF-1R\textsuperscript{37}, which play a key role in renal fibrosis\textsuperscript{38,39}.

We found that TEAD4 and FOXA1 are the top enriched master regulators of our top three suggested hubs VDR, RXRA, and IL-8. TEAD4, a member of TEF family, is a TF that plays a key role in the Hippo signaling pathway, a conserved pathway involved in cell proliferation, organ size, and tissue regeneration\textsuperscript{40}. Recent studies showed that Hippo signaling could regulate renal tubulointerstitial fibrosis mediated by several factors including TEAD family\textsuperscript{41}. In the present study, we suggest TEAD4 for the 1\textsuperscript{st} time as a key molecule in developing MN through regulation of some mediators of inflammation and immunity.

FOXA1 belongs to a superfamily of forkhead TFs that recently has been identified in T cells which, after ectopic expression, confers suppressive properties in a newly identified T-reg cell population\textsuperscript{42}. Involvement of FOXA1 in renal disease was suggested by Behr et al., whose experiment showed that its mutation leads to a defect in renal water homeostasis in vivo and nephrogenic diabetes in a mouse model\textsuperscript{43}. It is suggested here that impaired function of FOXA1 through aberrant regulation of regulatory T cells is most probably mediated by VDR, RXRA, and IL-8, which result in chronic inflammation and T-cell activation injury to the kidney during MN progression.

Downstream effector analysis indicated that two of our highly significant suggested hubs (i.e., VDR and RXRA) mostly regulate channel (e.g., KCNN3, KCNJ4, and CACNB3), receptor (e.g., LRP1, ERBB3, and RARA), and cellular cycle (e.g., CDKN1A and CCND2) regulatory genes (Fig. 2). This could be indicative of the importance of these target genes in the pathogenesis of MN. A short discussion on each of these effectors has been presented in Supporting Information file 2.

From the above discussion, we could draw the conclusion that 41 genes are considered as signature for MN, of which four genes (i.e., VDR, RXRA, IL-8, and SH3D2A) have priority based on network parameters. Analyses of the merged network from 50 significant
Figure 3. The abstract subnetwork of membranous nephropathy. The rectangles represent key master regulators, ellipses represent the most important hub genes, and the octagons represent downstream target genes. The directed edges show direction of regulation.

In conclusion, we highlight the role of two hub genes as they are TF themselves and have at least nine target genes in common (i.e., LRP1, ITGA3, KCNN3, KCNJ4, CACNB3, CDKN1A, CCND2, ERBB3, and RARA) and two common master regulators (i.e., TEAD4 and FOXA1). Accordingly, TEAD4 and FOXA1 are suggested as potential pharmaceutical targets, while VDR and RXRA are suggested as novel diagnostic candidate biomarkers for MN. Experimental validation of these candidate biomarkers in future studies can prove their application in the clinic. Furthermore, the assay of these candidate biomarkers in the biological fluids including blood and urine is promising for the non-invasive diagnosis of MN. Finally, we suggest the specific subnetwork for MN for the 1st time using a PPI approach (Fig. 3). The importance of such bioinformatics analyses on disease status is to find the potential genes that might be important in the pathogenesis of the disease when their changes may not be detected in the experiments because of low limit detection. Therefore, experiments focused on the optimization of the procedures to detect these target genes and molecules will elucidate their actual role in MN pathogenesis and may be useful in clinical diagnosis and targeted therapy. Accordingly, in silico approaches can pave the way for diagnosis and treatment and could thus be useful in clinical practice.

ACKNOWLEDGMENTS

The authors would like to thank the Chronic Kidney Disease Research Center (CKDRC) and Urology and Nephrology Research Center (UNRC), Shahid Beheshti University of Medical Sciences, Tehran - Iran, for their help and support. S. Kalantari would like to thank Nasibeh Khayer for the helpful comments.

Supplementary data

Supplementary data are available at Revista de Investigación Clínica online (www.clinicalandtranslational-investigation.com). These data are provided by the corresponding author and published online for the benefit of the reader. The contents of supplementary data are the sole responsibility of the authors.

REFERENCES