*Article*

# Identification of driver genes in renal stress condition using network clustering approach

**Mohd Murshad Ahmed**[1], **Safia Tazyeen**[1], **Rafat Ali**[1], **Aftab Alam**[1], **Md Zubbair Malik**[2], **Romana Ishrat**[1]
[1]Centre for Interdisciplinary Research in Basic Sciences, Jamia Millia Islamia, New Delhi 110025, India
[2]School of Computational and Integrative Sciences, Jawaharlal Nehru University, New Delhi 1100067, India
E-mail: rishrat@jmi.ac.in

*Received 25 February 2021; Accepted 30 March 2021; Published 1 September 2021*


**Abstract**
Chronic kidney disease (CKD) is a non-contagious, ageing-related and covert disease. It is a (chronic) disease of the kidneys leading to renal failure, huge public health problem worldwide and common comorbidity with type 2 diabetes mellitus (T2DM). Its presence and severity influence disease prognosis significantly. Identification of driver genes which regulate the CKD network is one of the main challenges in understanding its biological significance. We have analyzed microarray dataset and compare the gene expression profile of the patient with healthy control. Besides, we studied the gene regulatory networks that may help to understand the molecular mechanism in CKD. Further, after a comparative analysis of CKD and DKD. We proposed five driver genes, namely ALB, WT1, IL7R, PTPRC and DOCK2, that play an essential role in the pathogenesis of CKD and could serve as biomarkers. In the present study, we have mapped and analyzed the interactions of these five genes in the form of network and in addition to this we also tracked down the other essential, i.e. driver genes responsible for the modular nature of the network. The proposed study is based on network analysis approaches to predict some unknown CKD-associated genes, which can be validated as reliable candidates for further *in vitro*/*in vivo* experiment.

**Keywords** CKD; microarray data analysis; DEGs; driver gene; DKD.

**Network Biology**
**ISSN 2220-8879**
**URL: http://www.iaees.org/publications/journals/nb/online-version.asp**
**RSS: http://www.iaees.org/publications/journals/nb/rss.xml**
**E-mail: networkbiology@iaees.org**
**Editor-in-Chief: WenJun Zhang**
**Publisher: International Academy of Ecology and Environmental Sciences**

## 1 Introduction

Chronic kidney disease (CKD) is chronic, non- contagious disease and age factor kidney function alter disease that affects the population around the world (De Nicola and Zoccali, 2016). The protracted effect may become the higher risk of cardiovascular diseases (CVD) and finally leads to death. CKD prevalence higher with age. A higher prevalence of CKD can be seen in older age groups as compared to the young ones and is also dependent on gender (higher prevalence in females). The maximum burden of healthcare is in primary stages, around 35% affecting those over 70 years (Hill et al., 2016). As per the Kidney Disease

Improving Global Outcomes guidelines, CKD is defined as a set of abnormalities related to kidney's function/structure and is present for more than three months, with implications for health. It can be divided into five stages based on the glomerular filtration rate (GFR)(Levey et al., 2005). Early stages can be manifested with minor kidney damage which is usually marked by the presence of albumin in the urine test. An increase in levels of albumin is followed by a gradual decrease in renal function and eventually leading to the advanced stage (last stage) known as ESRD. It requires expensive renal replacement the rapy in the form of dialysis or transplantation (Stevens et al., 2010). ESRD is a significant health concern mostly in developed countries, with an annual growth rate of dialysis program somewhere in between 6% to 12% in over the past two decades. Our of all five stages, stage 3 CKD, has a high worldwide prevalence of around 11% - 13 % (Hill et al., 2016). Since symptoms always appear in later stages as GFR worsens (unclear statement). This has caused CKD to reach up to epidemic proportions. 10%–12% of the population, which makes up ~50% of the elderly population shows signs of kidney dysfunction, a condition associated with high morbidity and mortality (unclear)(Hill et al., 2016).

Research on CKD suggested the involvement of multiple factors when it comes to defining its stages. These factors include susceptibility factors (increasing vulnerability to kidney damage), initiation factors (which cause kidney damage), and progression factors (which cause worsening damage) (Levey et al., 2007). CKD is regarded as an independent risk factor for CVD events which is the primary cause of morbidity and mortality (Kellum, Lameire, and for the KDIGO AKI Guideline Work Group, 2013). There is a graded inverse relationship between CVD risk and glomerular filtration rate (GFR) which is independent of age, sex and other risk factors (Jha et al., 2013). It is described by analysis of kidney damage, protein uria (commonly using albumin to creatinine ratio (ACR)) and decreased kidney functions (below thresholds of GFR estimated from serum creatinine concentration).

The present work tries to establish an unbiased catalogue of change (up and down-regulation) in gene expression for the CKD and DKD samples using networks theoretical approach. We bestowed a 22 sample of CKD analysis of gene-expression from normal persons and patients representing an ethnically diverse population. We've gotten genes from our previous comparative analysis. These genes were found involved in the basic functioning of the cell, such as the rearrangement of the cytoskeleton, tissues development and activation of the immune system. The present analysis will be useful for the researchers to collect evidence against CKD working in in-vitro. Various gene and miRNA involve in diabatic nephropathy related pathogenesis. There have been various studies on CKD considering the feedback mechanisms in the CKD network, incorporating diabetic nephropathy. A recent report suggests a potential unexplored function of these proteins in CKD, both as predictive and therapeutic target biomarker. The study gives equal significance to the motifs and modules of the network to identify the driver gene and regulatory pathways not restricting only to overrepresented motifs establishing, identification, a relationship between them in CKD gene-disease association studies using network theoretical approach. It takes a holistic approach for predicting driver key disease genes and their pathways within network theoretical framework using datasets of CKD patients.

## 2 Materials and Methods

The detailed workflow of the CKD network and analysis is as follows (Fig. 1).

### 2.1 Dataset and processing

A set of 3 microarray datasets were selected from the NCBI Gene Expression Omnibus database (http://www.ncbi.nlm.nih.gov/geo/). The datasets include GSE70528, GSE11045 and GSE30122. According to the corresponding correlation between the probe and gene from the GPL570 platform HGU 133 Plus 2 (Affymetrix Human Genome U 133 plus 2.0 Array, Affymetrix, Santa Clara, CA, USA). The probe numbers

of the expression profile were converted into the corresponding gene symbols. One gene corresponds with multiple probes. Thus each gene has more than one expression value. Therefore, the average value was calculated and selected as the only representative value. A total of 54675 genes were involved in pre-processed data. Background correction and quartile data normalization were performed by the robust multi-array (RMA) average in R affypackage (Affymetrix). In the present study, we used LIMMA package that is a highly recommended procedure to measure the DEGs. It calculates F-test and t-test by applying the *Empirical Bayes* approach and reducing the standard errors, which give us steady and reproducible outcomes even with a less quantity of arrays. In pre-processing, convolution background was corrected, missing values were estimated, expression values were log2 transformed and then normalized. Both data CEL files pre-process under the MAS5 method. The expression value after summarization is further used for DEG's. The genes with p-value, less than 0.05 and fold change greater than 2 were selected as differentially expressed genes (DEGs) between normal and CKD samples.

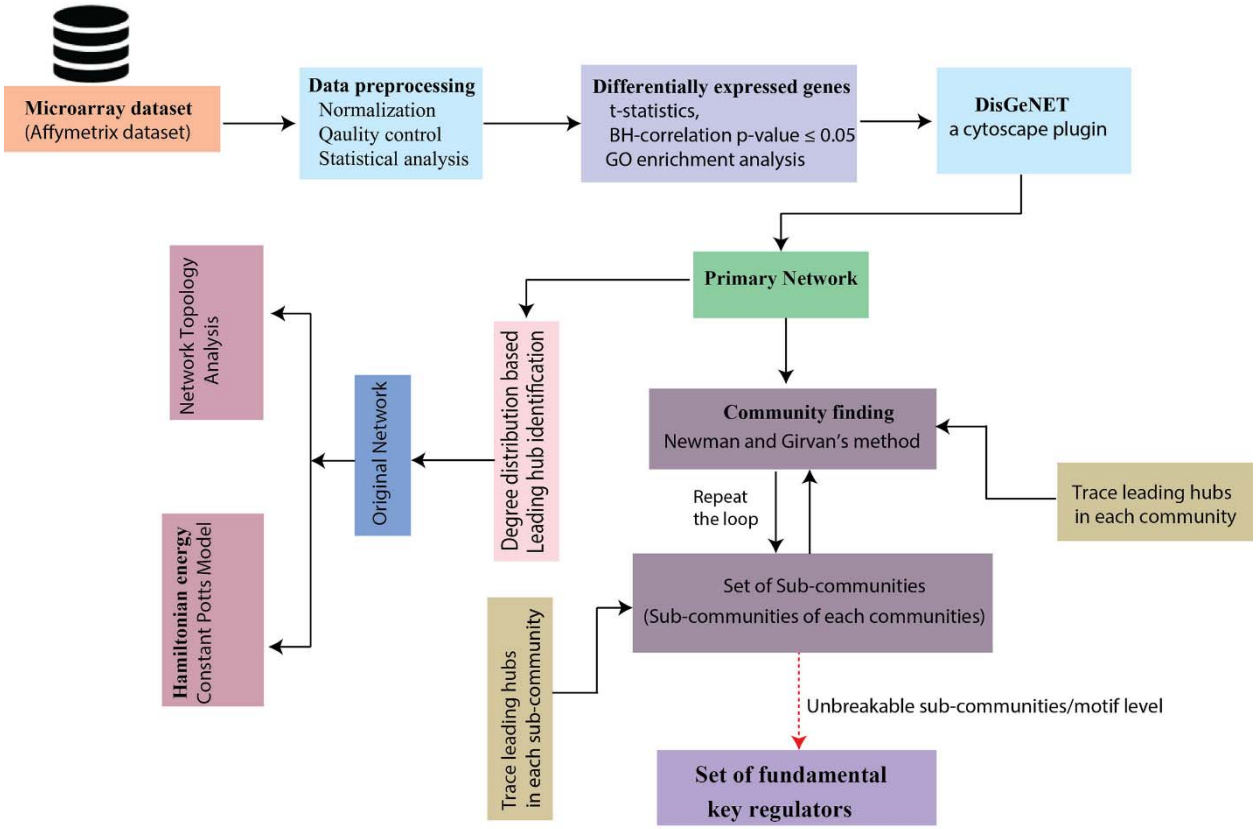


**Fig. 1** Schematic diagram of the workflow of the methods implemented in the study of CKD network.

## 2.2 Gene ontology (GO) enrichment analysis

Gene ontology terms give us the controlled vocabulary of terms divided into three categories like molecular function, biological processes and cellular components (Xin and Zhang, 2020). Thus, for a preliminary investigation into the functional differences of DEGs extracted were submitted to DAVID (Database for Annotation, Visualization and Integrated Discovery)an online software http://david.abcc.ncifcrf.gov/home.jsp)

to enrich the given set of DEGs to possible GO terms (Ashburner et al. 2000). A Benjamini-Hochberg (BH) p-value of less than 0.05 was used as a cutoff to select the significantly enriched pathway.

**2.3 Network construction**

The network was constructed using DisGeNET, an App available for Cytoscape versions 3.0 or higher. It allows the visualization gene-disease association's (GDAs) as bipartite graphs and additionally provides disease or gene-centric views of the data. It integrates data from curated databases with the information available from the literature. We provided DisGeNET with a list of DEG's obtained from our previous study.

**2.4 Method to the detection of levels of the organization and to identify driver gene**

Finding communities play a vital role in characterizing and detecting the nature and topological properties of the hierarchical network. Thus, describing the behaviour of the network at a different level of the hierarchy and accessing the organizing principle of the network. Leading Eigenvector (LEV) is one of the methods available for community detection that is reliable (in our case) as it calculates the eigenvalue for each edge, giving importance to links (edges), not nodes. Therefore, we apply LEV detection method in R using the '*igraph*' package. We used this method to detect modules from the main network, sub-modules from modules at each level of organization, and so on until we got only motifs (i.e. 3 nodes and 3 edges)(Sadi et al., 2010). For this process, we stick to the criterion of identifying any sub-module as the community by the presence of at least 1 motif (defined by $G(3,3)$).

**2.5 Calculation of Hamiltonian energy: Constant Pots Model**

To understand the compactness and organization of network at some state/level of the network, we calculated the Hamiltonian energy of that gene at that level. HE provides an understanding of the stability at global as well as community level (Ravasz, 2002). HE of network or any module can be calculated as in equation (1),

$$HE = \sum_i [e_c - \gamma n_c^2] \quad (1)$$

where, $e_c$ is the edges in the community or module, $n_c$ is the number of nodes in the same community or module and $\gamma$ being the resolution parameter which is set to 0.5 .

**3 Result and Discussion**

This work provides information in multiple microarray datasets on the structure and function of interacting genes. Our target was to find the essential genes by comparing the datasets that are related to gene functions and pathways. Total19 $cel$ files are selected from GSE70528 data out of which eight represents hypertension data, seven for CKD data and four for $hemodialysis$ data. Total 54,675 genes from $GSE70528$ and 22,278 genes from $GSE30122$ were filtered at $log_2(FC) = 6$ & 2 respectively. After comparing the expression of set 1 ($HCH$ and $GLOMERULI$ diabetes genes), we found 21 Up-regulated genes, while for the $set2$ containing (HCH and TUBULI diabetes genes) 8 Up-regulated genes were found. For the same comparison, some 36 and 32 down-regulated genes were found respectively from $set1$ and $set2$. Finally, we got five genes, of which four were up-regulated and six down-regulated (irrespective for their involvement in cancer). Identification of DEGs between the normal and treated conditions can be calculated in terms of the hypothesis (equation (2)).

$$H_0^i : \mu_N^i - \mu_T^i = 0 \,,\; H_1^i : \mu_N^i - \mu_T^i \neq 0 \quad (2)$$

The distribution of sample $\overline{X}_N^i - \overline{X}_T^i$ is used to test the hypothesis$H_0^i$, where $\overline{X}_N^i$ and $\overline{X}_T^i$ are the means of the (array) replicates in the control group and replicates in the treatment group, respectively, and $s_{iN}^2$ and $s_{iT}^2$ are the sample variances respectively. P-values were calculated for each gene to find the significance of genes see

**Table 1** UP and DOWN regulated genes showing statistics.

| Probe ids affymetrix | Gene | Fold Change | SD treat | SD wild | t-value | p-value |
|---|---|---|---|---|---|---|
| 211298_s_at | **ALB** | **9.65** | 0.012285 | 1.810692 | 23.245 | 1.82E-16 |
| 206952_at | **G6PC** | **6.94** | 0.064545 | 1.43397 | 20.968 | 1.45E-15 |
| 206067_s_at | **WT1** | **7.37** | 0.014634 | 1.570643 | 20.473 | 2.35E-15 |
| 213894_at | **THSD7A** | **6.70** | 0.113128 | 1.941436 | 14.893 | 1.23E-12 |
| 204891_s_at | LCK | -7.85 | 0.760188 | 0.500862 | -17.323 | 6.49E-14 |
| 205798_at | IL7R | -8.33 | 0.982075 | 0.445559 | -14.475 | 2.13E-12 |
| 207238_s_at | PTPRC | -6.24 | 0.203114 | 0.640783 | -33.195 | 1.24E-19 |
| 213160_at | DOCK2 | -7.48 | 0.777429 | 0.315255 | -16.463 | 1.76E-13 |
| 206025_s_at | TNFAIP6 | -6.70 | 0.860337 | 1.206504 | -11.797 | 9.96E-11 |
| 210044_s_at | LYL1 | -6.52 | 1.092371 | 0.76229 | -9.964 | 2.06E-09 |

To get an idea of what could be the effect of DEGs can only be possible when we have some preliminary insight into their function as the term gene ontology (GO) enrichment allows us to get the information about the gene. The driver genes were found to display the differential expression irrespective of the condition, state of disease and pattern of their expression (Zhang and Horvath, 2005). The common extracted DEGs total of 10 genes obtained divided into two sets. The one set of up-regulated genes are ALB (Albumin), G6PC (Glucose-6-Phosphatase Catalytic Subunit), WT1 (Wilms tumor 1), and THSD7A (thrombospondin type-1 domain-containing 7A). In contrast, another set of down-regulated genes are DOCK2 (Dedicator of cytokinesis 2), IL7R (Interleukin 7 Receptor), LCK (lymphocyte-specific protein tyrosine kinase), PTPRC (Protein Tyrosine Phosphatase, Receptor Type C), TNFAIP6 (Tumor Necrosis Factor-Inducible Gene 6 Protein) and LYL1 (lymphoblastic leukaemia derived sequence 1). These up-regulator genes and down-regulator genes are actively participating in cancer, and rearrangement of few genes are the most common genetic defect associated with T cell acute lymphoblastic leukaemia. Hence, we used DAVID, an online tool for GO enrichment and function for the list of genes is shown in Table 2.

### 3.1 Delineation of global network topological properties

The topological properties can access the structure of the network. These properties signify the association of nodes to their neighbours and elucidate their behavioural relationship between them.

The topology of the network defined by the probability of degree distributions $(P(k))$, clustering coefficient $(C(k))$ and neighbourhood connectivity $(C_N(k))$ exhibit power law. Following the accord of power-law, one can get the picture of whether there is the existence of scale freeness or not. So, the network constructed for CKD, by using the genes extracted from the microarray experiments follows the power-law distribution. The probability of $(P(k))$ degree distributions, $(C(k))$ clustering co-efficient and $(C_N(k))$ neighbourhood connectivity exhibit power-law or fractal nature (Fig. 2) and for complete network, it is given by topological properties of CKD network (Barabási and Oltvai, 2004).

Furthermore, we also detected the modules in a constructed network. The biological functional analysis finding of drugs and kinases by the by X2K shown in Table 3.

**Table 2** 10 genes showing BP, CC, MF by DAVID.

| GENE | GOTERM_BP_DIRECT | GOTERM_CC_DIRECT | GOTERM_MF_DIRECT |
|---|---|---|---|
| LCK | **GO:0006882** cellular zinc ion homeostasis | pericentriolar material | non-membrane spanning protein tyrosine kinase activity |
| TNFA1P6 | **GO:0007155** cell adhesion | extracellular space | protein binding |
| WT1 | **GO:0006355** regulation of transcription | Nucleus | nucleic acid binding |
| ALB | **GO:0001895** retina homeostasis | extracellular space | DNA binding |
| DOCK2 | **GO:0001766** membrane raft polarization | Intracellular | guanyl-nucleotide exchange factor activity |
| G6PC | **GO:0005980** glycogen catabolic process | integral component of endoplasmic reticulum membrane | glucose-6-phosphatase activity |
| IL7R | **GO:0000902** cell morphogenesis | external side of plasma membrane | cytokine receptor activity |
| PTPRC | **GO:0050852** T cell receptor signaling pathway | integral component of membrane | protein tyrosine phosphatase activity |
| THSD7A | **GO:0001525** angiogenesis | plasma membrane | tube formation in angiogenesis. |
| LYL1 | **GO:0001955** blood vessel maturation | Nucleus | RNA polymerase II regulatory region sequence-specific DNA binding |

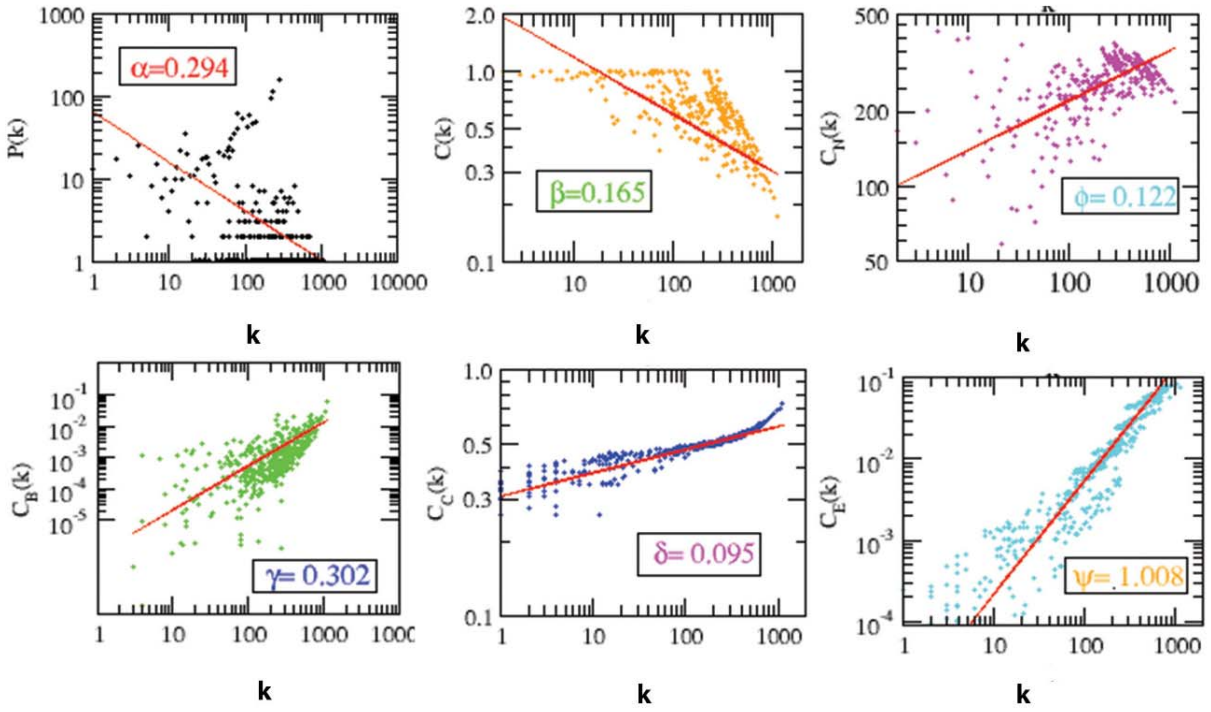


**Fig. 2** Topological properties of network $P(k)$, $C(k)$, $C_N(k)$, $C_B(k)$, $C_C(k)$ and $C_E(k)$ as a function of degree $k$.

The behaviour of the network is characterized by given equations (3)-(5), indicates hierarchy in synergy with the scale-free nature in the network. Fitting power law to the points in the topological properties for the network was done and verified by a standard statistical fitting procedure, where all p-values statistics is calculated for all the data sets against 2500 random samplings which were found to be larger than a critical value 0.1, and goodness of fits would be less than or equal to 0.33.

$$P(k) \sim k^{-\alpha} \tag{3}$$

$$C(k) \sim k^{-\beta} \tag{4}$$

$$C_N(k) \sim k^{\emptyset} \tag{5}$$

The positive value of $\emptyset$ for power law of neighborhood connectivity shows assertive nature of the network. While, the negative value of α for $P(k)$ shows the availability of each node in the network. The negative value in $\beta$ of $C(k)$ parameter shows disassociation in the communication between the nodes in the network (Albert and Barabási. 2002).

The basic centrality parameters, namely, betweenness $(C_B)$, closeness $(C_C)$ and eigen-vector $(C_E)$ centralities of the network also exhibit hierarchical behavior shown in (Fig. 2) given in following equation

$$\begin{bmatrix} C_B \\ C_C \\ C_E \end{bmatrix} = \begin{bmatrix} \gamma \\ \delta \\ \psi \end{bmatrix} = \begin{bmatrix} 0.302 \\ 0.095 \\ 1.008 \end{bmatrix} \tag{6}$$

The positive values of the centrality exponents indicates a strong regulating behavior of the nodes in the network (Barabasi, 2009).

**3.2 Finding the working driver genes in the network**

To understand the importance of a particular cellular function, one needs to identify fundamental molecular components (i.e. Genes). The interaction between all these regulating elements yields accountability to the network. The most traditional and basic approach to find this is to get the gist of formation of community in the network. The simple procedure described by the Neumann (Neumann, 2008), we find out the communities to prevails up to motif level for some to level 3 other to level 4 (Fig. 3).

Since the network is a follower of hierarchy, the communities have the capability of the network regulation with the hubs. The removal of some of them enhances the regulating abilities of the modules than the hubs, that is the direct picture of being strongly linked network.

The effectiveness of most connected nodes (i.e. hubs) changes with their regulating mechanisms. All of the hubs may not behave as the driver genes for clinical and drug target genes. To prove these, we traced through the network communities and found interesting results. However, few of the hubs (apart from the genes used for network building) can be significant, which we term as purely driver genes. These driver genes are defined as deeply rooted genes which can able to reach from primary network to motif level the regulating unit through different levels of organization via sub-communities Fig. 3. These regulators work at the level where the only effect visible is in terms of essential technologies and generally form the backbone of the network stability. They could become essential to the networks for both structural and functional integrity at different levels. It also works as an essential organizer in maintaining network stabilization whenever the network is under

external stress. These regulatory elements play the leading role in propagating the information and also work as receivers. These regulators serve as interconnections of far and near nodes in the network that may segregate even though they are physically far away from one another. As long as these fundamental or driver genes are there in the network deeply rooted in through a large number of levels of the organization, the network will have strong capability of defending any stress, and revert the changes provided by it Fig. 4A.

**Table 3** Drugs, Kinase, Pathways and TFs for 10 genes by X2K.

| GENE | TFs | KINASE | DRUGS | PATHWAYS BY KEGG |
|---|---|---|---|---|
| LCK | ETS2 | MAPK1 | tolmetin-4088 | T_Cell_Receptor_Signaling_Pathway |
| TNFA1P6 | ETS1 | MAPK14 | genistein-2695 | Galactose_Metabolism |
| WT1 | PEA3 | MAPK8 | acebutolol-6631 | Glycolysis_And_Gluconeogenesis |
| ALB | E4BP4 | MAPK3 | tretinoin-6170 | Adipocytokine_Signaling_Pathway |
| DOCK2 | CREBP1 | CSNK2A1 | prednicarbate-5119 | Starch_And_Sucrose_Metabolism |
| G6PC | AML | MAP2K3 | pentetrazol-1408 | Hematopoietic_Cell_Lineage |
| IL7R | HNF1 | JAK2 | ursodeoxycholic acid-7243 | Natural_Killer_Cell_Mediated_Cytotoxicity |
| PTPRC | ELK1 | PRKCD | valproic acid-23 | Cell_Adhesion_Molecules |
| THSD7A | STAT6 | MAP2K1 | rofecoxib-371 | Insulin_Signaling_Pathway |
| LYL1 | GATA | CSNK2A2 | verapamil-6287 | Jak_Stat_Signaling_Pathway |

We were able to identify some regulators out of large set of hubs in the CKD network, which are ALB, G6PC, WT1, THSD7A and DOCK2, PTPRC, LCK, LYL1, IL7R, TNFAIP6 respectively. These hubs regulate the process at different levels of network and motif. We calculated modularity averages of community at all level by *igraph* and make the modularity graph (Fig. 4B).

We further calculated the HE of the respective driver genes (see Methods) to unveil the importance of these genes in providing effective communication in the network up to motif level. The trend HE for these genes show (Fig. 4C) that they effectively participate in the organization of network that has a direct correlation to communication (Traag et al., 2011).

## 4 Conclusion

Gene expression analysis using microarray datasets from NCBI has given the opportunity to reasonably curate the set of important genes. We bolted out and reported a set of ten genes that are expressed differentially in various conditions inside kidney. Finally, the tracking down of these driver genes, we found that five genes, namely ALB, WT1, IL7R, PTPRC, DOCK2 were present even at the basic motif level of the network. Interestingly, it is seen that PTPRC and IL7R were found closely associated with each other and form the same motif. While genes ALB, DOCK2 and WT1 diverge to independent motifs. Thus, it can be said the expression disorders associated with the pair of PTPRC and IL7R genes can be of prime focus to the researcher. This network also follows the modular hierarchical nature, which represents low degree DEGs are much modular than higher degree DEGs. Due to cohesively group of genes are functioning for the conclusive response of particular function,

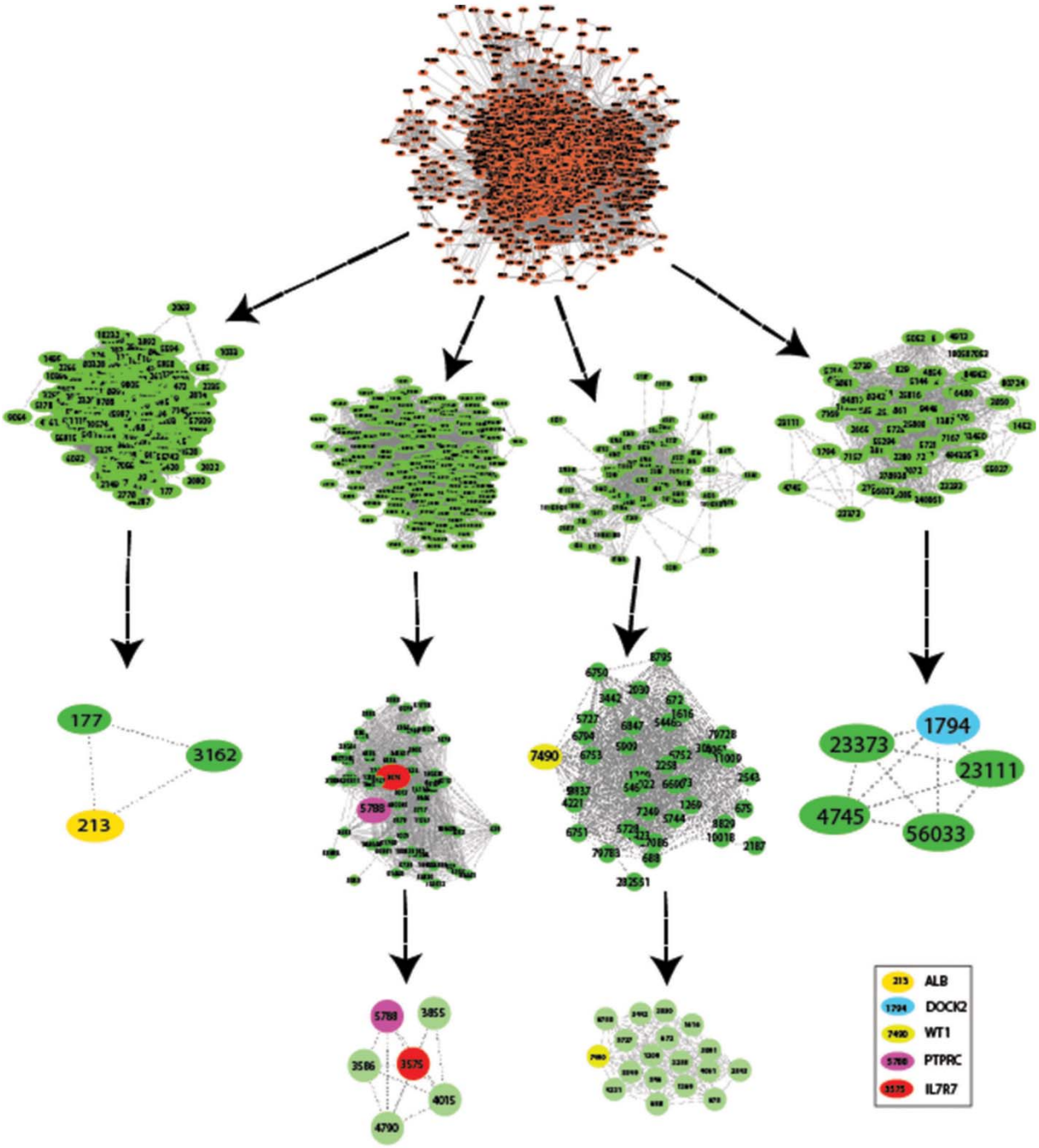


**Fig. 3** Tracking down the presence of the driver genes within different modules at different level of the network.

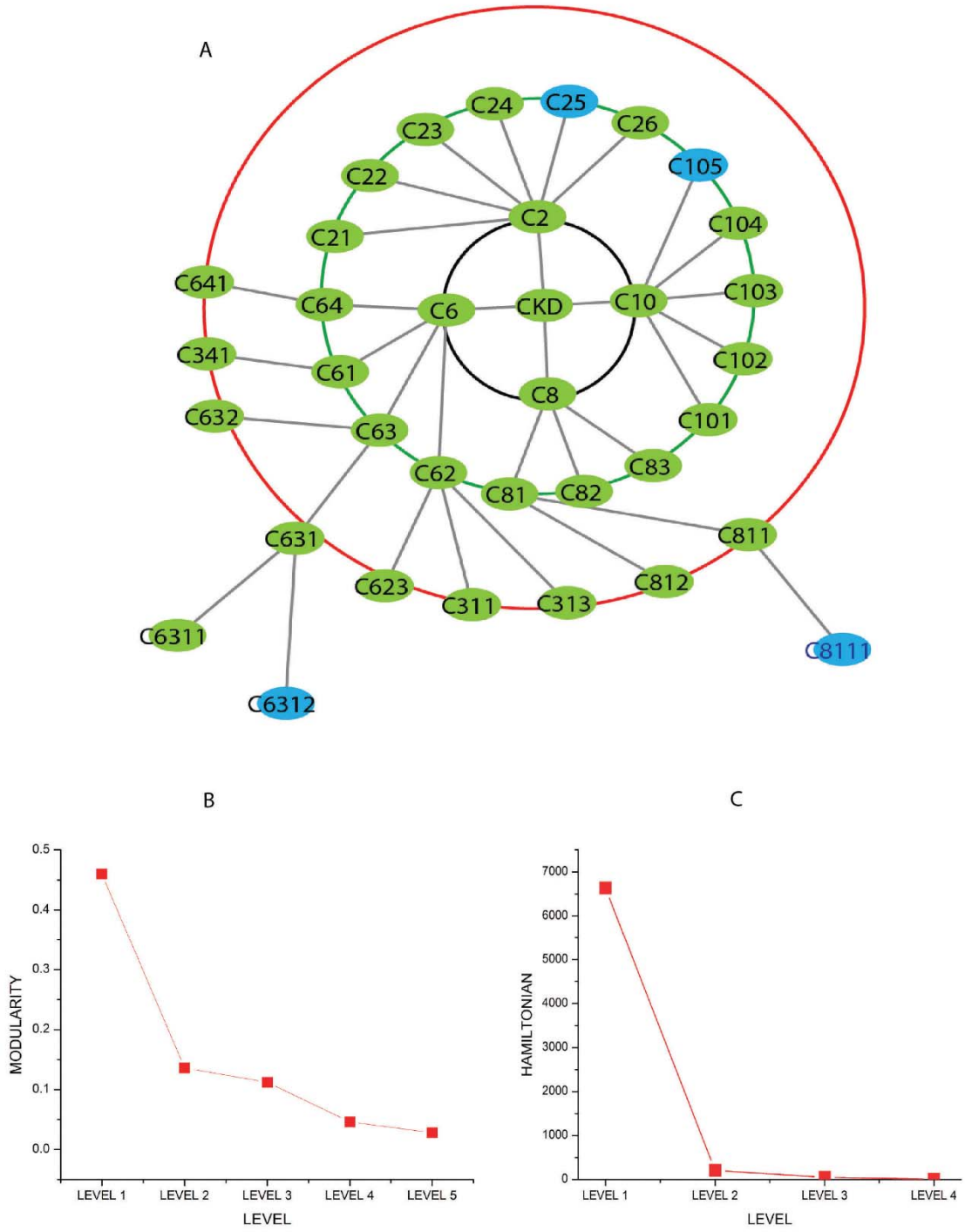


**Fig. 4 A** community diagram showing level of gene tracing **B:** modularity graph of level versus modularity **C:** Hamiltonian Energy graph.

## Acknowledgments

MMA and ST financially supported by Indian Council of Medical Research, Government of India under Senior Research Fellow. FTS No. 525507. MZM and AA financially supported by Department of Health and Research, Ministry of Health and Family Welfare, Government of India under young scientist. MMA, MZM and RI conceptualized the work, MMA MZM did Data curation, MMA. RA and MZM preparation of the

figures. MMA, RA, MZM and RI wrote the manuscript. MMA, RA, ST, AA, MZM authors read and approved the manuscript.

**References**

Abdel-Daim MM, Abdou RH. 2015. Protective effects of diallyl sulfide and curcumin separately against thallium-induced toxicity in rats. Cell Journal, 17(2): 379

Abdel-Daim MM, Khalifa HA, Abushouk AI, Dkhil MA, Al-Quraishy SA. 2017. Diosmin attenuates methotrexate-induced hepatic, renal, and cardiac injury: a biochemical and histopathological study in mice. Oxidative Medicine And Cellular Longevity, 2017: 3281670

Abdeen A, Abdelkader A, Abdo M, Wareth G, Aboubakr M, Aleya L, Abdel-Daim M. 2019. Protective effect of cinnamon against acetaminophen-mediated cellular damage and apoptosis in renal tissue. Environmental Science and Pollution Research, 26(1): 240-249

Abdel-Daim MM, Dessouki AA, Abdel-Rahman HG, Eltaysh R, Alkahtani, S. 2019. Hepatorenal protective effects of taurine and N-acetylcysteine against fipronil-induced injuries: the antioxidant status and apoptotic markers expression in rats. Science of the Total Environment, 650: 2063-2073

Abdel-Daim MM, Abdellatief SA. 2018. Attenuating effects of caffeic acid phenethyl ester and betaine on abamectin-induced hepatotoxicity and nephrotoxicity. Environmental Science and Pollution Research, 25(16): 15909-15917

Ahmed MM, et al. 2017. In silico based analysis of CKD expressions data in correlation with diabetes mellitus unveils biomarker gene. Journal of Theoretical Computer Science, 4: 156

Albert R, Barabási AL. 2002. Statistical mechanics of complex networks. Reviews of Modern Physics, 74(1): 47

Ashburner M, Ball CAC, Blake JA, et al. 2000. Gene Ontology: tool for the unification of biology. Nature Genetics, 25(1): 25-29

Baer R. 1993. TAL1, TAL2 and LYL1: a family of basic helix-loop-helix proteins implicated in T cell acute leukaemia. Seminars in Cancer Biology, 4(6): 341-347

Barabási AL. 2009. Scale-free networks: a decade and beyond. Science, 325(5939): 412-413

Barabasi AL, Oltvai ZN. 2004. Network biology: understanding the cell's functional organization. Nature Reviews Genetics, 5(2): 101

Barrett T, Troup DB, Wilhite SE, Ledoux P, Evangelista C, et al. 2010. NCBI GEO: archive for functional genomics data sets-10 years on. Nucleic Acids Research, 39: 1005-1010

Berriz GF, Roth FP. 2008. The synergizer service for translating gene, protein and other biological identifiers. Bioinformatics, 24(19): 2272-2273

Couser WG, Remuzzi G, Mendis S, Tonelli M. 2011. The contribution of chronic kidney disease to the global burden of major noncommunicable diseases. Kidney International, 80(12): 1258-1270

Daim MM, Abd Eldaim MA, Hassan AG. 2014. Trigonella foenum-graecum ameliorates acrylamide-induced toxicity in rats: Roles of oxidative stress, proinflammatory cytokines, and DNA damage. Biochemistry and Cell Biology, 93(3): 192-198

De Nicola L, Zoccali C. 2015. Chronic kidney disease prevalence in the general population: heterogeneity and concerns. Nephrology Dialysis Transplantation, 31(3): 331-335

Dkhil MA, Kassab RB, Al-Quraishy S, Abdel-Daim MM, Zrieq R., Moneim AEA. 2018. *Ziziphus spinachristi* (L.) leaf extract alleviates myocardial and renal dysfunction associated with sepsis in mice. Biomedicine and Pharmacotherapy, 102: 64-75

Fawzy MS, Abu AlSel BT, Al Ageeli E, Al-Qahtani SA, Abdel-Daim MM, Toraih EA. 2018. Long non-coding RNA MALAT1 and microRNA-499a expression profiles in diabetic ESRD patients undergoing dialysis: a preliminary cross-sectional analysis. Archives of Physiology and Biochemistry, 1-11

Gautier L, Cope L, Bolstad BM, Irizarry RA. 2004. affy-analysis of Affymetrix GeneChip data at the probe level. Bioinformatics, 20: 307-315

Gu DF, Shi YL, Chen YM, Liu HM, Ding YN, Liu XY, Yuan ZY. 2013. Prevalence of chronic kidney disease and prediabetes and associated risk factors: a community-based screening in Zhuhai, Southern China. Chinese Medical Journal, 126(7): 1213-1219

Hill NR, Fatoba ST, Oke JL, Hirst JA, O'Callaghan CA, Lasserson DS, Hobbs FR. 2016. Global prevalence of chronic kidney disease–a systematic review and meta-analysis. PloS one, 11(7): e0158765

Ibrahim A, Al-Hizab FA, Abushouk AI, Abdel-Daim MM. 2018. Nephroprotective effects of benzyl isothiocyanate and resveratrol against cisplatin-induced oxidative stress and inflammation. Frontiers in Pharmacology, 9

Ibrahim A, Eldaim MAA, Abdel-Daim MM. 2016. Nephroprotective effect of bee honey and royal jelly against subchronic cisplatin toxicity in rats. Cytotechnology, 68(4): 1039-1048

Jha V, Garcia-Garcia G, Iseki K, Li Z, Naicker S, Plattner B, Yang CW. 2013. Chronic kidney disease: global dimension and perspectives. The Lancet, 382(9888): 260-272

Kellum JA, Lameire N, KDIGO AKI Guideline Work Group. 2013. Diagnosis, evaluation, and management of acute kidney injury: a KDIGO summary (Part 1). Critical Care, 17(1): 204

Levey AS, Atkins R, Coresh J, Cohen EP, Collins AJ, Eckardt KU, Powe NR. 2007. Chronic kidney disease as a global public health problem: approaches and initiatives–a position statement from Kidney Disease Improving Global Outcomes. Kidney International, 72(3): 247-259

Levey AS, Eckardt, KU, Tsukamoto Y, Levin A, Coresh J, Rossert J, et al. 2005. Definition and classification of chronic kidney disease: a position statement from Kidney Disease: Improving Global Outcomes (KDIGO). Kidney International, 67(6): 2089-2100

Malik MZ, Alam MJ, Ishrat R, Agarwal SM, Singh RB. 2017. Control of apoptosis by SMAR 1. Molecular BioSystems, 13(2): 350-362

Malik MZ, Chirom K, Ali S, Ishrat R, Somvanshi P, Singh RB. 2019. Methodology of predicting novel key regulators in ovarian cancer network: a network theoretical approach. BMC Cancer, 19(1): 1129

Mangangcha IR, Malik MZ, Küçük Ö, Ali S, Singh RB. 2019. Identification of key regulators in Prostate cancer from gene expression datasets of patients. Scientific Reports, 9: 16420

Neumann ID. 2008. Brain oxytocin: a key regulator of emotional and social behaviours in both females and males. Journal of Neuroendocrinology, 20(6): 858-865

Pinero J, Queralt-Rosinach N, Bravo A, Deu-Pons J, Bauer-Mehren A, Baron M, et al. 2015. DisGeNET: a discovery platform for the dynamical exploration of human diseases and their genes. Database, 2015

Sadi S, Ögüdücü Ş, Uyar AŞ. 2010. An efficient community detection method using parallel clique-finding ants. In: IEEE Congress on Evolutionary Computation. IEEE

Rachidi SM, Qin T, Sun S, Zheng WJ, Li Z. 2013. Molecular profiling of multiple human cancers defines an inflammatory cancer-associated molecular pattern and uncovers KPNA2 as a uniform poor prognostic cancer marker. PloS one, 8(3): e57911.

Ravasz E, Somera AL, Mongru DA, Oltvai ZN, Barabási AL. 2002. Hierarchical organization of modularity in metabolic networks. Science, 297(5586): 1551-1555

Singh SS, Haobijam D, Malik MZ, Ishrat R, Singh RB. 2018. Fractal rules in brain networks: Signatures of self-organization. Journal of Theoretical Biology, 437: 58-66

Smyth GK. 2004. Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. Statistical Applications in Genetics and Molecular Biology, 3: 3

Stevens LA, Viswanathan G, Weine, DE. 2010. Chronic kidney disease and end-stage renal disease in the elderly population: current prevalence, future projections, and clinical significance. Advances in Chronic Kidney Disease, 17(4): 293-301

Traag VA, Van Dooren P, Nesterov Y. 2011. Narrow scope for resolution-limit-free community detection. Physical Review E, 84(1): 016114

Xin SH, Zhang WJ. 2020. Construction and analysis of the protein protein interaction network for the olfactory system of the silkworm *Bombyx mori*. Archives of Insect Biochemistry and Physiology, 100(1): e21737

Zhang B, Horvath S. 2005. A general framework for weighted gene co-expression network analysis. Statistical Applications in Genetics and Molecular Biology, 4(1): Article 17