ANRIL Regulates Multiple Molecules of Pathogenetic Significance in Diabetic Nephropathy

Parisa Sooshtari
Western University
Biao Feng
Western University
Saumik Biswas
Western University
Michael Levy
Western University
Hanxin Lin
Western University
Zhaoliang Su
Jiangsu University
Subrata Chakrabarti (subrata.chakrabarti@lhsc.on.ca)
Western University

Research Article

Keywords: Hyperglycemia, aberrant synthesis, ANRIL knockout, streptozotocin

Posted Date: June 2nd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-563215/v1

ANRIL regulates multiple molecules of pathogenetic significance in diabetic nephropathy

Parisa Sooshtari[1,2], Biao Feng[1], Saumik Biswas[1], Michael Levy[1], Hanxin Lin[1,3], Zhaoliang Su[4], Subrata Chakrabarti[1,3]

[1]Department of Pathology and Laboratory Medicine, Western University and [2]Ontario Institute for Cancer Research, [3]London Health Sciences Centre, London, Ontario, Canada and Dept. of Immunology, [4]Jiangsu University, Jiangsu, PR China.

Address Correspondence to:

Subrata Chakrabarti
Department of Pathology and Laboratory Medicine
Western University
London, Ontario, N6A 5C1, Canada
DSB 4033
Tel: (519) 685-8500 X36350
Fax: (519) 663-2930
Email: subrata.chakrabarti@lhsc.on.ca

Running title: ANRIL in Diabetic Nephropathy.

Word count: Abstract: 252 Text (excluding abstract): 3213

Number of figures: 7 (Suppl. Fig. 2)

Number of Tables: 3 (Suppl. Table 4)

Abstract:

Background: Hyperglycemia-induced transcriptional alterations lead to aberrant synthesis of a large number of pathogenetic molecules leading to functional and structural damage to multiple end organs including the kidneys. Diabetic nephropathy (DN) remains a major cause of end stage renal disease. Multiple epigenetic mechanisms, including alteration of long non-coding RNAs (lncRNAs) may play a significant role mediating the cellular transcriptional activities. We have previously shown that lncRNA ANRIL may mediate diabetes associated molecular, functional and structural abnormalities in DN. Here we explored downstream mechanisms of ANRIL alteration in DN.

Methods: We used renal cortical tissues from ANRIL knockout (KO) mice and wild type (WT) B6 mice, with or without streptozotocin (STZ) induced diabetes for RNA sequencing. The differentially expressed genes were identified using edgeR and DESeq2 computational methods. KEGG and Reactome pathway analyses and network analyses using STRING and IPA were subsequently performed.

Results: Diabetic animals showed hyperglycemia, reduced body weight gain, polyuria and increased urinary albumin. Both albuminuria and polyuria were corrected in the KO diabetic mice. RNA analyses showed Diabetes induced alterations of a large number of transcripts in the wild type (WT) animals. ANRIL knockout (KO) prevented a large number of such alterations. The altered transcripts include metabolic pathways, apoptosis, extracellular matrix protein synthesis and degradation, NFKB related pathways, AGE-RAGE interaction pathways etc. ANRIL KO prevented majority of these pathways.

Conclusion: These findings suggest that as ANRIL regulates a large number of molecules of pathogenetic significance, it may potentially be a drug target for DN and other chronic diabetic complications.

Introduction:

Nephropathy is one of the debilitating chronic complications of diabetes. Diabetic nephropathy remains a major cause of end stage renal disease (ESRD) [1]. Starting with glomerular hyperfiltration there is progressive renal damage, manifesting as albuminuria, declining glomerular filtration rate [GFR], and ultimately ESRD. At the microscopic level the kidneys show glomerular hypertrophy, glomerulosclerosis, tubulointerstitial inflammation and fibrosis. Although over the last few years various novel therapies have been developed, a large number of diabetics still progress towards ESRD [2]. Hence better understanding of the pathogenesis and further delineation of pathogenetic mechanisms are needed to develop biomarkers for early diagnosis and novel treatment targets.

Hyperglycemia has established itself as the key factor for the initiation and progression of all chronic diabetic complications including DN [3]. However, in a tissue specific way the effect of hyperglycemia may vary, causing alteration of specific gene expression and ultimately causing tissue damage [4,5]. Multiple mechanisms contribute to the initiation and progression of DN. Hyperglycemia leads to multiple cellular structural and functional changes in the organs. These include hemodynamic changes and cellular metabolic changes leading to altered gene transcription of several such vasoactive and growth factors such as VEGF, TGF β, setting the stage for tissue damage and organ failure [5]. Microscopically, kidneys demonstrate glomerular basement membrane thickening, increased mesangial matrix deposition, eventually manifested as glomerulosclerosis and interstitial fibrosis [6]. At the biochemical level, there is deposition and reduced degradation of extracellular matrix (ECM) proteins such as fibronectin [FN],

type IV collagen (Col1α4) etc [6]. Functionally, there is increased urinary albumin excretion ultimately leading to renal failure [2,3].

Epigenetic alterations regulating the synthetic process of macromolecules possibly plays a major role in all chronic diseases including DN [7,8]. Such modifications include DNA methylation, histone modifications through acetylation and methylation and alterations of noncoding RNAs (ncRNAs) [9]. ncRNAs consists of several classes of RNA molecules including microRNAs and long noncoding RNAs (lncRNAs). Of particular relevance to this research, lncRNAs (200-2000nt) lack protein-coding capacity and are capable of regulating gene expression at both transcriptional and translational levels [10]. They can regulate local(*cis*) and distal[*trans]* genes by a large number of mechanisms and play key regulatory roles in diverse biological processes [10,11].

lncRNA ANRIL, situated in the antisense direction of p15/CDKN2B-p16/CDKN2A-p14/ARF (*INK4b-ARF-INK4a*) gene cluster, has 19 exons and spans126.3kb [11]. We have recently reported that ANRIL is upregulated and plays a pathogenetic roles in several diabetic complications including DN [12,13]. We have further shown that it works through an interaction with EZH2 of polycomb repressive complex 2(PRC2) and histone acetylator, p300 [12,13]. ANRIL locus is a hotspot for multiple disease-associated polymorphisms and DNA alterations and has been steadily associated with cardiovascular diseases, cancer, diabetes, glaucoma and other conditions [14]. We have also demonstrated oxidative DNA damage in diabetes is a potential mechanism of ANRIL upregulation in diabetes [15,16] In our previous study, we have shown that ANRIL blockade prevents diabetes-associated molecular, functional and structural abnormalities in the context of DN [13].

To gain a better understanding of the downstream mechanisms and pathways of action of ANRIL and to identify potential drug targets, we took RNA-sequencing and bioinformatics analysis-based approach. To this extent, we analyzed renal cortical tissues of a type 1 model of diabetes, with or without ANRIL knockout.

Methods:

Animals: All animals were cared for according to the guiding principles in the care and use of animals. The experiments were approved by Western University Animal Care and Veterinary Services. All experiments conform to the guide for care and use of laboratory animals published by the NIH (NIH publication no. 85-23, revised in 1996).

ANRIL knockout (129S6/SvEvTac-Del[4C4-C5]1Lap/Mmcd) mice, were obtained from Mutant Mouse Resource & Research Centre (MMRC, Davis, CA). In the mouse, ANRIL is located on Chromosome 4 within the *p15/CDKN2B-p16/CDKN2A-p14/ARF* gene cluster (INK4 locus) in an anti-sense manner. ANRILKO mice were generated on a 129S6/SvEv background where a 70kb region of the mouse genome corresponding to the human 58-kb non-coding RNA associated with coronary artery disease risk interval was deleted [15]. We have previously demonstrated lack of ANRIL expression in the ANRILKO mice tissues [12,13].

A chemically-induced mouse model of Type 1 diabetic was generated with intraperitoneally injected streptozotocin (five doses, 50mg/kg in citrate buffer, pH5.6) on consecutive days [17,18]. As our previous studies used male mice and to have a cost-containment, in this investigation we kept the same focus. We do however, understand the importance of using mice of both sexes and plan to carry out such analyses in the future. Age and sex-matched littermate controls received equal volume of citrate buffer. Diabetes was confirmed by blood glucose measurement (>20mmol/L) from tail vein using a glucometer. Animals were monitored for variations in body weight and blood glucose. Mice were sacrificed after 8 weeks of the onset of diabetes and tissues were collected.

Renal cortices were dissected out from the kidneys and kept frozen at 80^{o}C for further analysis. A total of four groups of mice [n=3/groups] were used, namely a) non-diabetic wild type control (WT-C), b) Wild type mice with diabetes (WT-D), c) ANRIL knockout control (KO-C) and d) ANRIL knockout mice with diabetes (KO-D)

RNA sequencing:

Three biological replicates were used for each of the four groups. Total RNA was extracted using the RNeasy Mini Kit (QIAGEN, Toronto, ON). The libraries were prepared by using the TruSeq RNA V2 kit. Sequencing was performed by Macrogen Corp. (Rockville, MD) on an Illumina sequencer with NovaSeq6000 S4 flow cell.

Bioinformatics analysis of RNA-sequencing data:

An average of 42 million (minimum of 27 million) paired-end reads were generated for each sample. Samples were aligned using STAR version 2.7.3a [19] to mouse genome mm10 with GENCODE M23 primary annotations. An average of 30 million reads aligned uniquely (minimum 20 million). Subsequent analysis was performed using R version 3.6.1. Count matrices were generated using the summarized overlaps function from the Genomic Alignments package version 1.20.1 [20]. Differential expression analysis was performed using DESeq2 version 1.24.0 [21] and edgeR version 3.28.0 24 [22]. A gene was considered to be differentially expressed if it had an adjusted p-value < 0.01.

KEGG and Reactome pathway analysis:

After we identified differentially expressed genes, we applied KEGG [23,24,25] and Reactome [26] pathway analyses to identify pathways that are enriched for significant

genes in (a) wild-type control vs. wild-type diabetic mice (WT-C vs. WT-D), and (b) wild-type diabetic vs. diabetic knockout mice (WT-D vs. KO-D). In order to perform KEGG and Reactome pathway analyse, we chose to use the built-in functions available in STRING database and web resource (https://string-db.org). We identified a list of significant pathways with false discovery rate (FDR) < 0.01. We also measured the number of down-regulated and up-regulated genes from our list of genes that map to each pathway. Additionally, STRING provided a list of significant biological processes and molecular functions enriched for each comparison at the FDR < 0.01.

Protein-protein interactions and cluster analysis:

In addition to employing STRING functions to identify KEGG and Reactome pathways and the relevant biological processes and molecular functions, we used STRING to visualize protein-protein interaction (PPI) networks for each comparison (WT-C vs. WT-D, and WT-D vs. KO-D), and to identify clusters of interacting genes. Briefly, SRTING integrates both predicted and known PPIs to predict functional interactions of proteins, where each gene/gene product is presented by a node in the network, and the biological relationship between pairs of genes is represented by an edge between the two nodes.

We used STRING implementation of Markov Cluster Algorithm (MCL) with the default parameters to identify clusters of interacting genes (i.e. sub-networks) in each PPI network. We then considered each major cluster individually and identified their enriched pathways and molecular functions. This provided insight into the underlying molecular mechanisms at a finer resolution (i.e. at the cluster-level).

IPA analysis:

To further identify molecular sub-networks associated with each comparison (WT-C vs. WT-D and WT-D vs. KO-D), we used QIAGEN Ingenuity Pathway Analysis (QIAGEN IPA) software (QIAGEN Inc., https://www.qiagenbio- informatics.com/products/ingenuity-pathway-analysis), as an alternative network analysis approach. We uploaded a list of gene symbols along with their corresponding edgeR adjusted P values and log fold changes into the software, and set the analysis parameter to include genes with FDR < 0.01. Briefly, IPA uses the Ingenuity Pathways Knowledge Base to identify networks of the genes based on their connectivity and ranks networks based on the assigned scores. These scores take into the account the size of the network and the number of genes to predict if a network is relevant. Once the networks are ranked based on their assigned scores, they are presented by graphs that indicate the molecular relationships between genes or gene products. Here genes are presented by nodes, and the biological relationships between pairs of genes are represented by edges.

Results

Diabetes induced metabolic alteration and renal damage are prevented by nullifying ANRIL production:

Diabetic animals were maintained for a period of eight weeks following onset of hyperglycemia and were compared with Age- and sex- matched controls. Wild type diabetic (WT-D) animals showed hyperglycemia, reduced body weight, polyuria (urine volume >25ml/day in WT-D, compared to <3 ml /day in WT-C) and glycosuria (not shown). These are characteristic metabolic features of poorly controlled type 1 diabetes. ANRIL

KO diabetic mice (KO-D) remained hyperglycemic. However, they showed reduction of urine volume (<10ml/day) compared to WT-D mice. In keeping with previous analyses, we also calculated urinary albumin creatinine ratio (Figure 1). Urinary albumin/creatinine ratios, were elevated in WT-D mice compared to wild type controls (WT-C), as a result of renal damage and were normalized in the KO-D mice, indicating a protective effect of nullifying ANRIL.

Diabetes-induced alterations of a large number of transcripts are prevented following ANRIL KO:

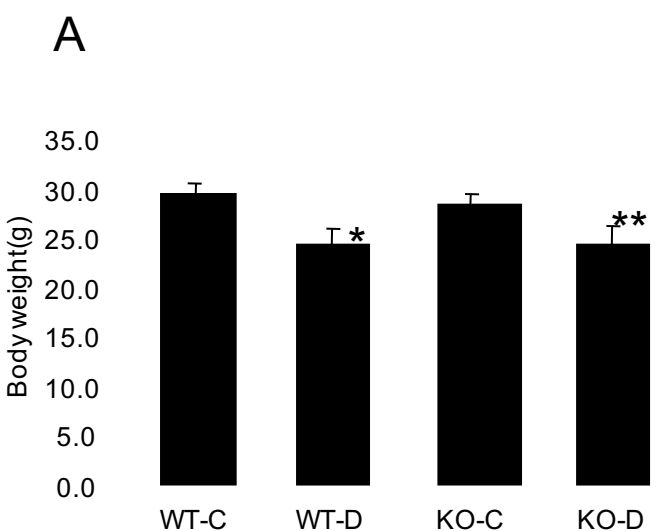
We used two bioinformatics programs, (edgeR and DESeq2) to identify the genes that are differentially expressed between multiple groups of samples. Our aims were first to identify the differentially expressed transcripts in the kidneys in the context of DN by comparing wild-type control (WT-C) with wild-type diabetic (WT-D) mice. We then tried to identify the differentially expressed transcripts that are regulated through ANRIL by comparing wild-type diabetic (WT-D) and diabetic knockout (KO-D) tissues. We also examined any alterations of transcripts in the kidneys of non-diabetic animals caused nullification of ANRIL by comparing WT-C and KO-C. Overall alterations of the transcripts are graphically depicted in Figure 2. Differentially expressed genes detected by EdgeR are mostly a subset of those detected by DESeq2. Since EdgeR was more stringent and may generate less false positive results, it is possibly a better option compared to DESeq2. However, in both such analyses, minimal number of alterations of transcripts were seen at the basal level. Specifically, EdgeR didn't detect any differentially expressed transcripts between wild type (WT-C) and non-diabetic knockout controls (KO-C). On the

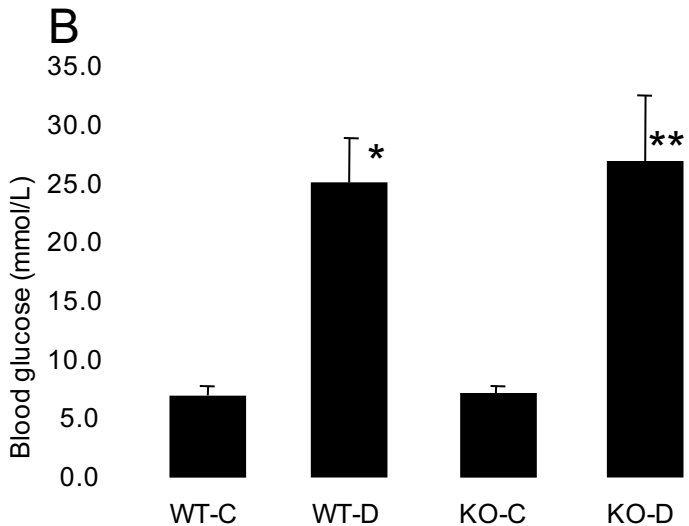
other hand, poorly controlled Diabetes caused alterations of a large number of transcripts in the wild type (WT-D) animals, while the number was significantly reduced in ANRIL knockout diabetic mice (KO-D). For example, in edgeR analyses, number of differentially expressed genes between WT-C and WT-D at p value < 0.05 was 4,198 and that at the p value <0.01 was 1,555. ANRIL knockout reduced these numbers to 1,059 and 175, for p values < 0.05 and 0.01, respectively (Figure 2).

ANRIL regulates multiple DN related transcripts:

We then sought to identify specific pathways that are altered in DN. We used KEGG and Reactome pathway analyses for those genes differentially expressed between WT-C and WT-D mice. Both KEGG and Reactome pathway analyses at P < 0.05 level showed a large number of alterations of transcripts (Suppl. Figure 1A and Suppl. Figure 2A). Of specific relevance to DN (WT-C vs. WT-D; P < 0.01), these transcripts include metabolic pathways, apoptosis, extracellular matrix protein synthesis and degradation, NFKB related pathways, AGE-RAGE interaction pathways etc. (Fig. 3A, 3E). ANRIL KO prevented majority of these pathways. At the P < 0.01 level, only group of transcripts which were not normalized in KEGG analysis were mRNAs related to metabolic pathways (Figures 3C). Reactome analysis did not reveal any residual uncorrected pathways. Given that these animals are hyperglycemic, the finding the transcripts in metabolic pathways in KEGG pathway analysis that were not normalized is not surprising (Figure 3). The detailed listings of these pathways and transcripts are in the Suppl. Table 1 (KEGG analysis; WT-C vs. WT-D), Suppl. Table 2 (KEGG analysis; WT-D vs. KO-D) and in Suppl. Table 3 (Reactome analysis; WT-C vs. WT-D).

Figure 1


A
Body weight(g)
35.0
30.0
25.0
20.0
15.0
10.0
5.0
0.0
*
**
WT-C
WT-D
KO-C
KO-D


B
Blood glucose (mmol/L)
35.0
30.0
25.0
20.0
15.0
10.0
5.0
0.0
*
**
WT-C
WT-D
KO-C
KO-D

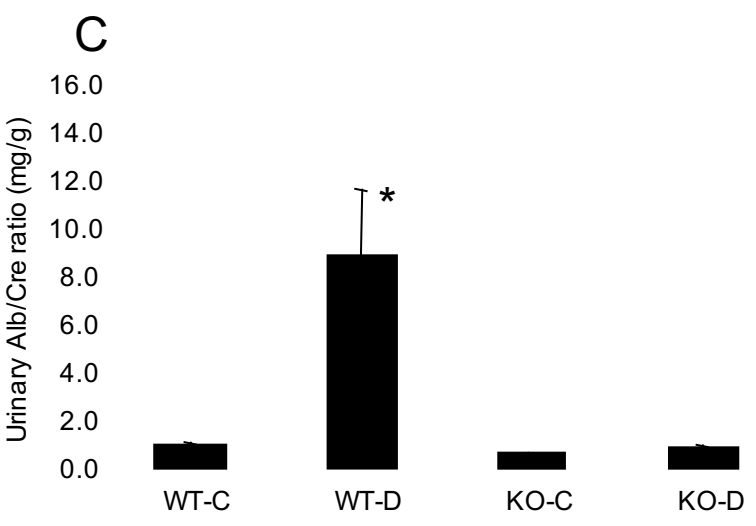

C
Urinary Alb/Cre ratio (mg/g)
16.0
14.0
12.0
10.0
8.0
6.0
4.0
2.0
0.0
*
WT-C
WT-D
KO-C
KO-D

Figure 2

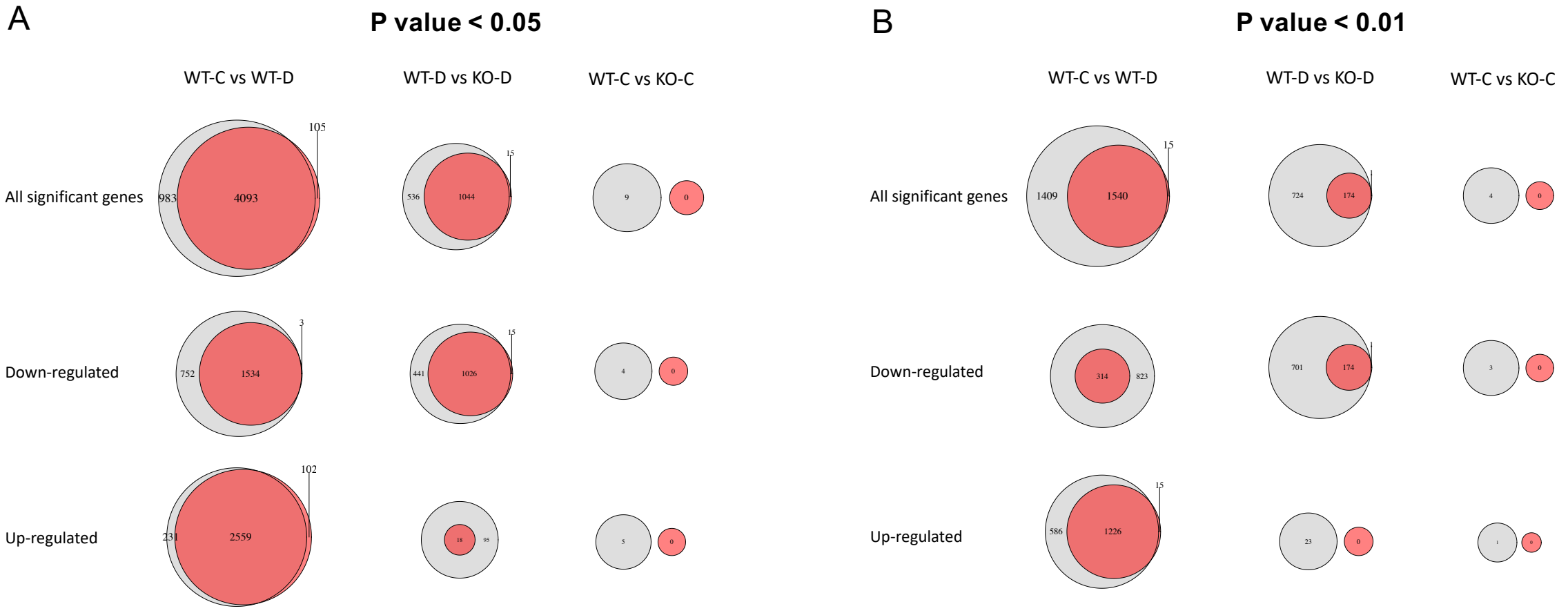
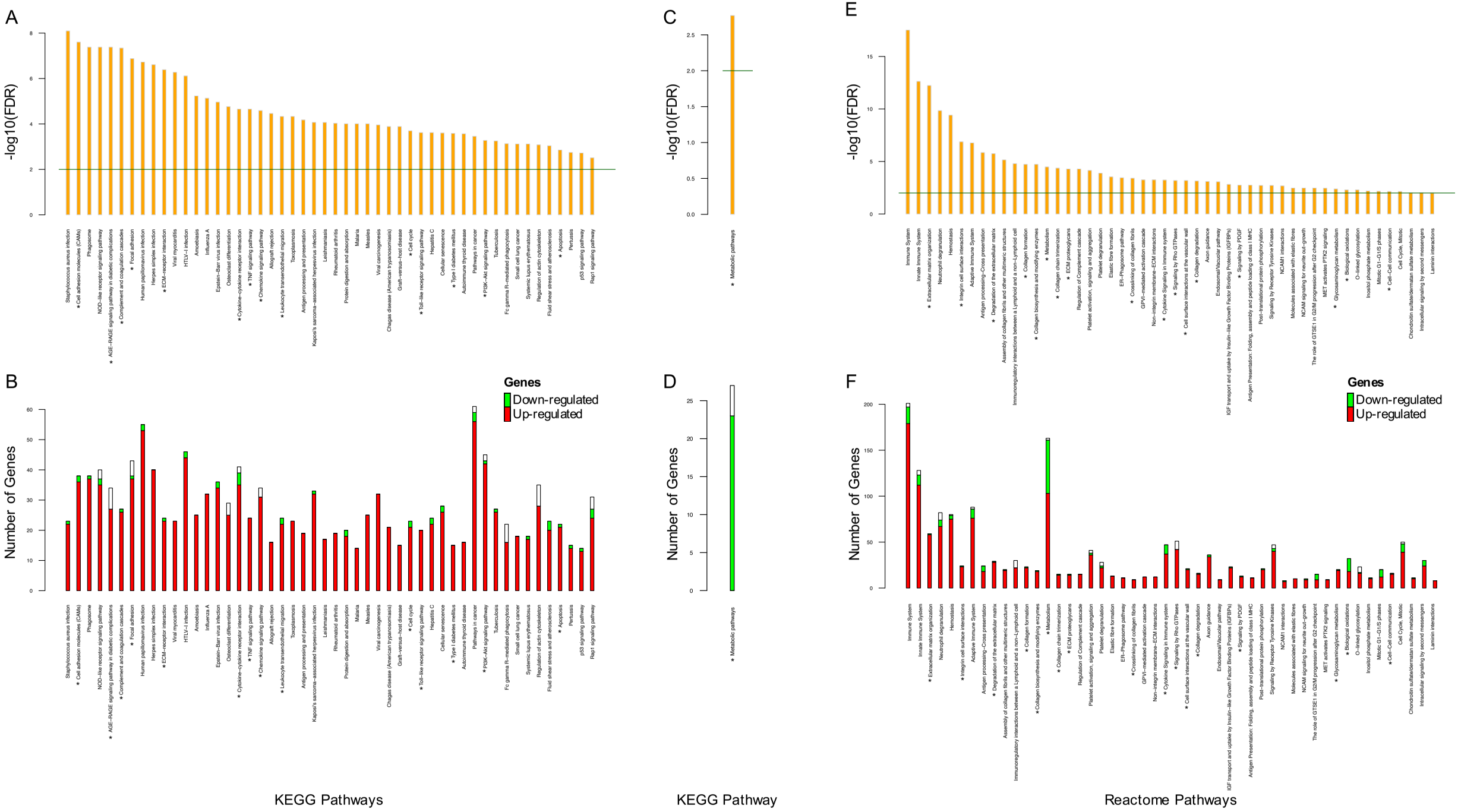

A
P value < 0.05
WT-C vs WT-D
WT-D vs KO-D
WT-C vs KO-C
All significant genes
983
4093
105
536
1044
15
9
0
Down-regulated
752
1534
3
441
1026
15
4
0
Up-regulated
231
2559
102
18
95
5
0
B
P value < 0.01
WT-C vs WT-D
WT-D vs KO-D
WT-C vs KO-C
All significant genes
1409
1540
15
724
174
1
4
0
Down-regulated
314
823
701
174
1
3
0
Up-regulated
586
1226
15
23
0
1
0

Figure 3


A
-log10(FDR)
Staphylococcus aureus infection
★ Cell adhesion molecules (CAMs)
Phagosome
NOD−like receptor signaling pathway
★ AGE−RAGE signaling pathway in diabetic complications
★ Complement and coagulation cascades
★ Focal adhesion
Human papillomavirus infection
Herpes simplex infection
★ ECM−receptor interaction
Viral myocarditis
HTLV−I infection
Amoebiasis
Influenza A
Epstein−Barr virus infection
Osteoclast differentiation
★ Cytokine−cytokine receptor interaction
★ TNF signaling pathway
★ Chemokine signaling pathway
Allograft rejection
★ Leukocyte transendothelial migration
Toxoplasmosis
Antigen processing and presentation
Kaposi's sarcoma−associated herpesvirus infection
Leishmaniasis
Rheumatoid arthritis
Protein digestion and absorption
Malaria
Measles
Viral carcinogenesis
Chagas disease (American trypanosomiasis)
Graft−versus−host disease
★ Cell cycle
★ Toll−like receptor signaling pathway
Hepatitis C
Cellular senescence
★ Type I diabetes mellitus
Autoimmune thyroid disease
Pathways in cancer
★ PI3K−Akt signaling pathway
Tuberculosis
Fc gamma R−mediated phagocytosis
Small cell lung cancer
Systemic lupus erythematosus
Regulation of actin cytoskeleton
Fluid shear stress and atherosclerosis
★ Apoptosis
Pertussis
p53 signaling pathway
Rap1 signaling pathway
B
Number of Genes
Genes
Down-regulated
Up-regulated
KEGG Pathways
C
-log10(FDR)
★ Metabolic pathways
D
Number of Genes
★ Metabolic pathways
KEGG Pathway
E
-log10(FDR)
Immune System
Innate Immune System
★ Extracellular matrix organization
Neutrophil degranulation
Hemostasis
★ Integrin cell surface interactions
Adaptive Immune System
Antigen processing−Cross presentation
★ Degradation of the extracellular matrix
Assembly of collagen fibrils and other multimeric structures
Immunoregulatory interactions between a Lymphoid and a non−Lymphoid cell
★ Collagen formation
★ Collagen biosynthesis and modifying enzymes
★ Metabolism
★ Collagen chain trimerization
★ ECM proteoglycans
Regulation of Complement cascade
Platelet activation, signaling and aggregation
Platelet degranulation
Elastic fibre formation
ER−Phagosome pathway
★ Crosslinking of collagen fibrils
GPVI−mediated activation cascade
Non−integrin membrane−ECM interactions
★ Cytokine Signaling in Immune system
★ Signaling by Rho GTPases
★ Cell surface interactions at the vascular wall
★ Collagen degradation
Axon guidance
Endosomal/Vacuolar pathway
IGF transport and uptake by Insulin−like Growth Factor Binding Proteins (IGFBPs)
★ Signaling by PDGF
Antigen Presentation: Folding, assembly and peptide loading of class I MHC
Post−translational protein phosphorylation
Signaling by Receptor Tyrosine Kinases
NCAM1 interactions
Molecules associated with elastic fibres
NCAM signaling for neurite out−growth
The role of GTSE1 in G2/M progression after G2 checkpoint
MET activates PTK2 signaling
★ Glycosaminoglycan metabolism
★ Biological oxidations
O−linked glycosylation
Inositol phosphate metabolism
Mitotic G1−G1/S phases
★ Cell−Cell communication
Cell Cycle, Mitotic
Chondroitin sulfate/dermatan sulfate metabolism
Intracellular signaling by second messengers
Laminin interactions
F
Number of Genes
Genes
Down-regulated
Up-regulated
Reactome Pathways

Overall, these data indicate that lncRNA ANRIL regulate multiple transcripts with pathogenetic role in DN. Furthermore, it is of particular interest that compared to the normal (WT) animal ANRIL KO mice without diabetes had no significant alterations with respect to mRNA expression.

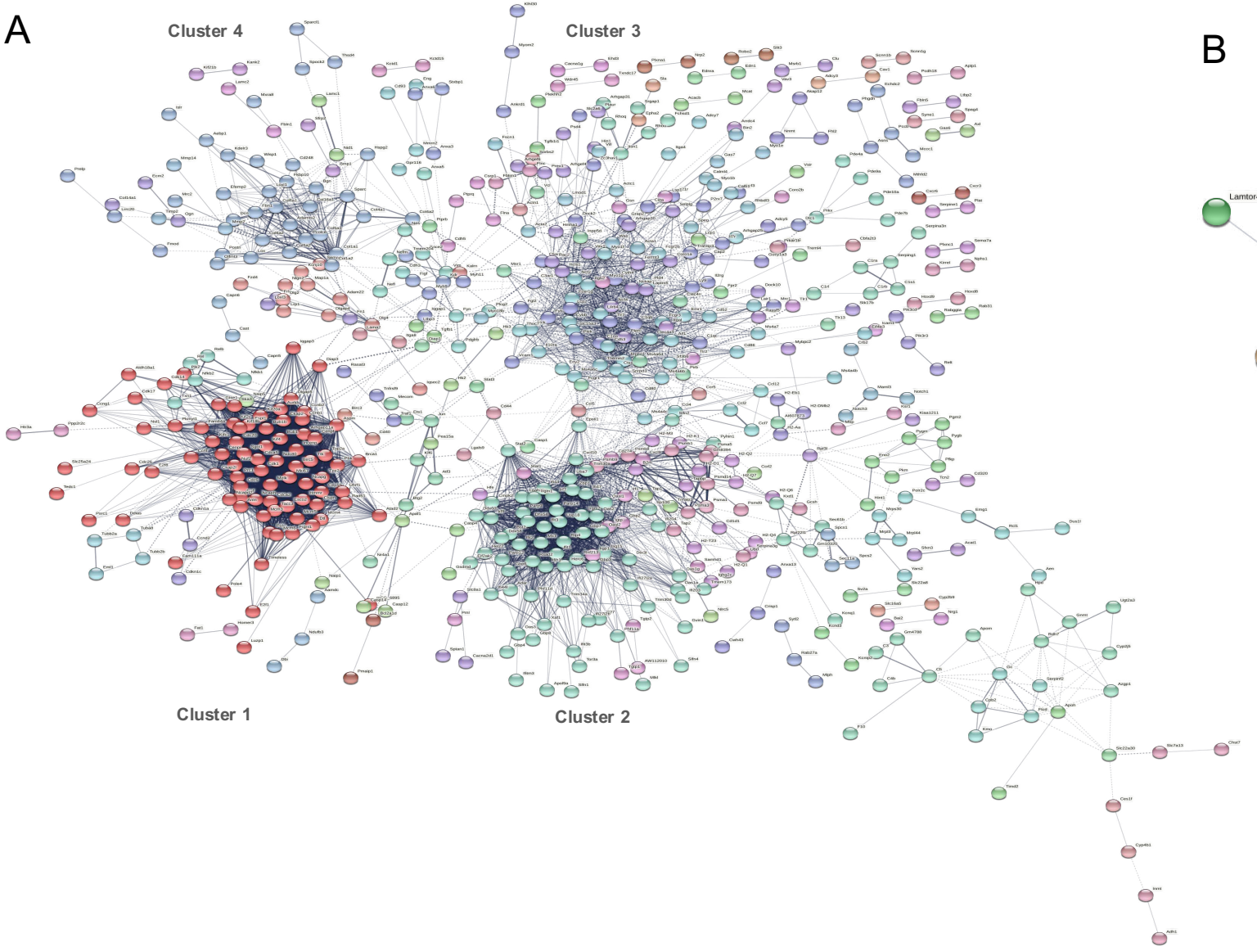
Protein-protein interactions and Cluster analysis:

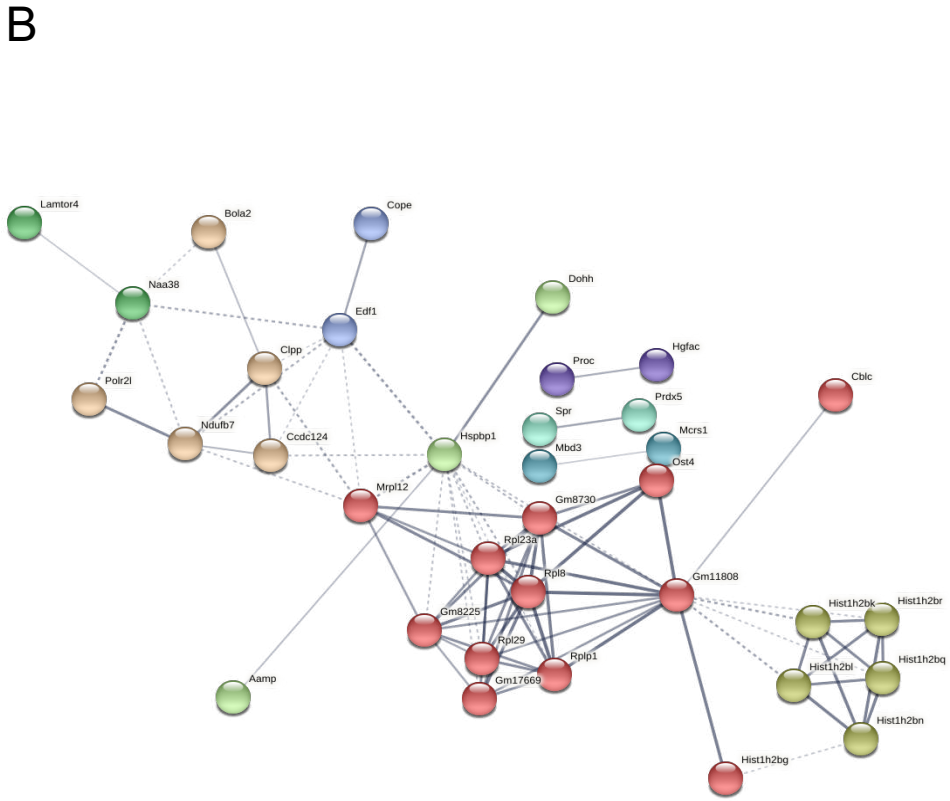
We performed protein-protein interaction (PPI) network analysis using STRING in order to uncover relationships between various differentially expressed transcripts (adjusted p value < 0.01). We observed that the transcripts that were differentially significant between wild-type diabetes and wild-type controls (WT-C vs. WT-D) formed multiple sub-networks (Figure 4A). These finding suggest that there are extensive interaction of multiple molecules and pathways, which works in a co-ordinated fashion to produce diabetes induced changes in the kidneys.

We further undertook a cluster analysis using MCL to identify the major clusters of differentially expressed transcripts in our PPI network. Such analyses revealed that four major clusters (with 29, 31, 70 and 81 genes per cluster) existed in the PPI network of the differentially expressed genes in the renal tissues in DN. In comparison, when we compared WT diabetic animals with KO diabetic animals (WT-D vs. KO-D) only smaller size clusters (with between 2 to 12 genes per cluster) were seen (Figure 4B). A diagrammatic representation of these clusters is shown in Figure 4.

We reasoned that each cluster of highly interacting genes in WT-C vs WT-D PPI network (Figure 4A) may have their own distinct functions, that might be obscured when we

Figure 4


A
Cluster 4
Cluster 3
Cluster 1
Cluster 2


B
Lamtor4
Bola2
Cope
Naa38
Edf1
Dohh
Clpp
Polr2l
Proc
Hgfac
Prdx5
Cblc
Spr
Ndufb7
Ccdc124
Hspbp1
Mcrs1
Mbd3
Ost4
Mrpl12
Gm8730
Rpl23a
Rpl8
Gm11808
Hist1h2bk
Hist1h2br
Gm8225
Rpl29
Rplp1
Hist1h2bl
Hist1h2bq
Aamp
Gm17669
Hist1h2bn
Hist1h2bg

consider all genes in the network together. We, therefore, identified KEGG and Reactome pathways, molecular functions and biological processes of each individual cluster (Suppl. Table 4). A detailed listing of the top KEGG pathways for each of the four major clusters is presented in Table 1. Some of the key genes and pathways altered in diabetes were identified in these clusters including collagens, TGFβ, PDGFR, PARPs, along with molecules which haven't been characterised yet in the context of DN.

Canonical pathway and sub-network analysis using IPA: We carried our further analyses using QIAGEN IPA software. Our analyses showed that a large number of transcripts in overlapping canonical pathways are altered in the kidneys in diabetes. Figures 5A and 5B depict overlaps of enriched IPA canonical pathways for wild-type control vs. wild-type diabetic mice (WT-C vs. WT-D); and wild-type diabetic vs. diabetic knockout mice (WT-D vs. KO-D), respectively. As Figure 5 suggests the canonical pathways in WT-C vs. WT-D are highly overlapping; while there are less interactions in the canonical pathways of WT-D vs. KO-D. This implies that the altered transcripts are normally expressed in tissues, in diabetes, they are overexpressed causing subsequent tissue damage and development of DN. ANRIL deletion corrected most of such alterations (Figure 5).

We further used QIAGEN IPA software to identify the top-ranked networks, and their relevant molecules and top diseases and functions for WT-C vs. WT-D (Table 2). Figure 6 shows networks 3, 4 and 17 of Table 2. These networks correspond to TGFβ1, VEGF and collagen molecules. We selected these networks as these are well established molecules of pathogenetic significance in DN [1,2,5,6]. Similar analysis of WT-D and KO-D showed only few sub-networks (Table 3). Figure 7 shows some of the known sub-

Table 1: STRING Protein-Protein Interaction Networks for WT-C vs. WT-D

A Cluster 1 (Red)

KEGG Pathway	Observed Gene Count	False Discovery Rate (FDR)	Matching Proteins
Cell cycle	16	2.60E-18	Cdc20,Cdk1,Mcm4,Pkmyt1,Mcm6,Bub1,Ccnb2,Bub1b,Mcm3,Espl1,Ttk,Ccnb1,Cdc26,Cdc6,E2f1,Mcm5
Oocyte meiosis	9	1.55E-08	Cdc20,Cdk1,Pkmyt1,Sgol1,Bub1,Ccnb2,Espl1,Ccnb1,Cdc26
DNA replication	6	1.37E-07	Pole,Mcm4,Mcm6,Mcm3,Pole4,Mcm5
Progesterone-mediated oocyte maturation	6	1.81E-05	Cdk1,Pkmyt1,Bub1,Ccnb2,Ccnb1,Cdc26
p53 signaling pathway	5	6.37E-05	Cdk1,Ccng1,Ccnb2,Ccnb1,Gtse1
HTLV-I infection	7	0.00059	Cdc20,Pole,Ccnb2,Bub1b,Cdc26,Pole4,E2f1
MicroRNAs in cancer	4	0.0134	Brca1,Ccng1,Cdca5,E2f1
Platinum drug resistance	3	0.0195	Brca1,Top2a,Birc5
Cellular senescence	4	0.0222	Cdk1,Ccnb2,Ccnb1,E2f1
Base excision repair	2	0.0371	Pole,Pole4
Homologous recombination	2	0.0459	Brca1,Rad51

B Cluster 2 (Aquamarine)

KEGG Pathway	Observed Gene Count	False Discovery Rate (FDR)	Matching Proteins
Influenza A	13	3.52E-13	Rsad2,Eif2ak2,Irf7,Ifih1,Oas3,Ddx58,Cxcl10,Oas2,Oas1a,Oas1g,Stat2,Adar,Irf9
Measles	11	1.27E-11	Eif2ak2,Irf7,Ifih1,Oas3,Ddx58,Oas2,Oas1a,Oas1g,Stat2,Adar,Irf9
Herpes simplex infection	12	3.10E-11	Eif2ak2,Irf7,Ifih1,Oas3,Ddx58,Oas2,Sp100,Oas1a,Oas1g,Ifit1,Stat2,Irf9
NOD-like receptor signaling pathway	11	5.80E-11	Irf7,Oas3,Gbp7,Oas2,Oas1a,Oas1g,Stat2,Gbp3,Ifi204,Tmem173,Irf9
Hepatitis C	10	1.53E-10	Eif2ak2,Irf7,Oas3,Ddx58,Oas2,Oas1a,Oas1g,Ifit1,Stat2,Irf9
RIG-I-like receptor signaling pathway	7	1.89E-08	Dhx58,Irf7,Ifih1,Ddx58,Cxcl10,Isg15,Tmem173
Cytosolic DNA-sensing pathway	6	0.000000287	Irf7,Zbp1,Ddx58,Cxcl10,Adar,Tmem173
Necroptosis	5	0.00061	Eif2ak2,Zbp1,Mlkl,Stat2,Irf9
Human papillomavirus infection	6	0.0024	Eif2ak2,Oasl1,Oasl2,Isg15,Stat2,Irf9
Hepatitis B	4	0.0035	Irf7,Ifih1,Ddx58,Stat2
Kaposi's sarcoma-associated herpesvirus infection	4	0.0104	Eif2ak2,Irf7,Stat2,Irf9
Viral carcinogenesis	4	0.0104	Eif2ak2,Irf7,Sp100,Irf9
TNF signaling pathway	3	0.0118	Cxcl10,Mlkl,Ifi47

C Cluster 3 (Sky Blue)

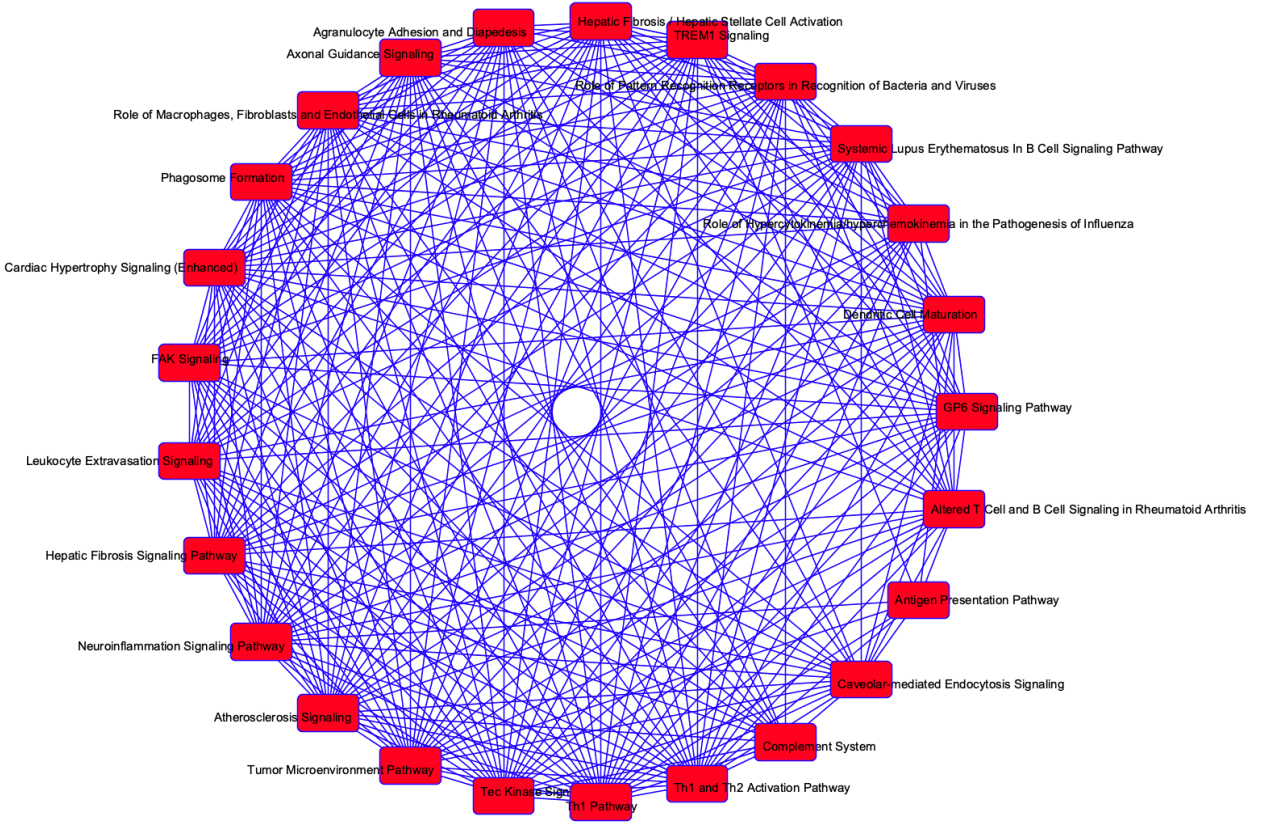
KEGG Pathway	Observed Gene Count	False Discovery Rate (FDR)	Matching Proteins
Leishmaniasis	5	4.28E-06	Itgb2,Cybb,Itga4,Fcgr3,Ncf2
Phagosome	5	1.80E-04	Itgb2,Cybb,Ctss,Fcgr3,Ncf2
Tuberculosis	5	1.80E-04	Itgb2,Ctss,Il10ra,Fcer1g,Fcgr3
Leukocyte transendothelial migration	4	5.50E-04	Itgb2,Cybb,Itga4,Ncf2
Osteoclast differentiation	3	1.38E-02	Lilrb4,Fcgr3,Ncf2
Cell adhesion molecules (CAMs)	3	0.0238	Itgb2,Cd86,Itga4
Intestinal immune network for IgA production	2	0.0238	Cd86,Itga4
Staphylococcus aureus infection	2	0.0294	Itgb2,Fcgr3

D Cluster 4 (Cornflower Blue)

KEGG Pathway	Observed Gene Count	False Discovery Rate (FDR)	Matching Proteins
Protein digestion and absorption	8	7.61E-12	Col6a1,Col6a2,Col1a1,Col5a1,Col1a2,Col4a1,Col4a2,Col5a2
ECM-receptor interaction	7	1.55E-10	Col6a1,Col6a2,Col1a1,Col1a2,Col4a1,Col4a2,Hspg2
Focal adhesion	6	1.19E-06	Col6a1,Col6a2,Col1a1,Col1a2,Col4a1,Col4a2
Human papillomavirus infection	6	1.94E-05	Col6a1,Col6a2,Col1a1,Col1a2,Col4a1,Col4a2
PI3K-Akt signaling pathway	6	1.96E-05	Col6a1,Col6a2,Col1a1,Col1a2,Col4a1,Col4a2
AGE-RAGE signaling pathway in diabetic complications	4	0.0000286	Col1a1,Col1a2,Col4a1,Col4a2
Amoebiasis	4	0.0000295	Col1a1,Col1a2,Col4a1,Col4a2
Relaxin signaling pathway	4	0.0000582	Col1a1,Col1a2,Col4a1,Col4a2
Proteoglycans in cancer	4	0.00026	Col1a1,Col1a2,Dcn,Hspg2
Small cell lung cancer	2	0.0115	Col4a1,Col4a2
Platelet activation	2	0.0178	Col1a1,Col1a2

Figure 5

A


Agranulocyte Adhesion and Diapedesis
Axonal Guidance Signaling
Role of Macrophages, Fibroblasts and Endothelial Cells in Rheumatoid Arthritis
Phagosome Formation
Cardiac Hypertrophy Signaling (Enhanced)
FAK Signaling
Leukocyte Extravasation Signaling
Hepatic Fibrosis Signaling Pathway
Neuroinflammation Signaling Pathway
Atherosclerosis Signaling
Tumor Microenvironment Pathway
Tec Kinase Signaling
Th1 Pathway
Hepatic Fibrosis / Hepatic Stellate Cell Activation
TREM1 Signaling
Role of Pattern Recognition Receptors in Recognition of Bacteria and Viruses
Systemic Lupus Erythematosus In B Cell Signaling Pathway
Role of Hypercytokinemia/hyperchemokinemia in the Pathogenesis of Influenza
Dendritic Cell Maturation
GP6 Signaling Pathway
Altered T Cell and B Cell Signaling in Rheumatoid Arthritis
Antigen Presentation Pathway
Caveolar-mediated Endocytosis Signaling
Complement System
Th1 and Th2 Activation Pathway

B

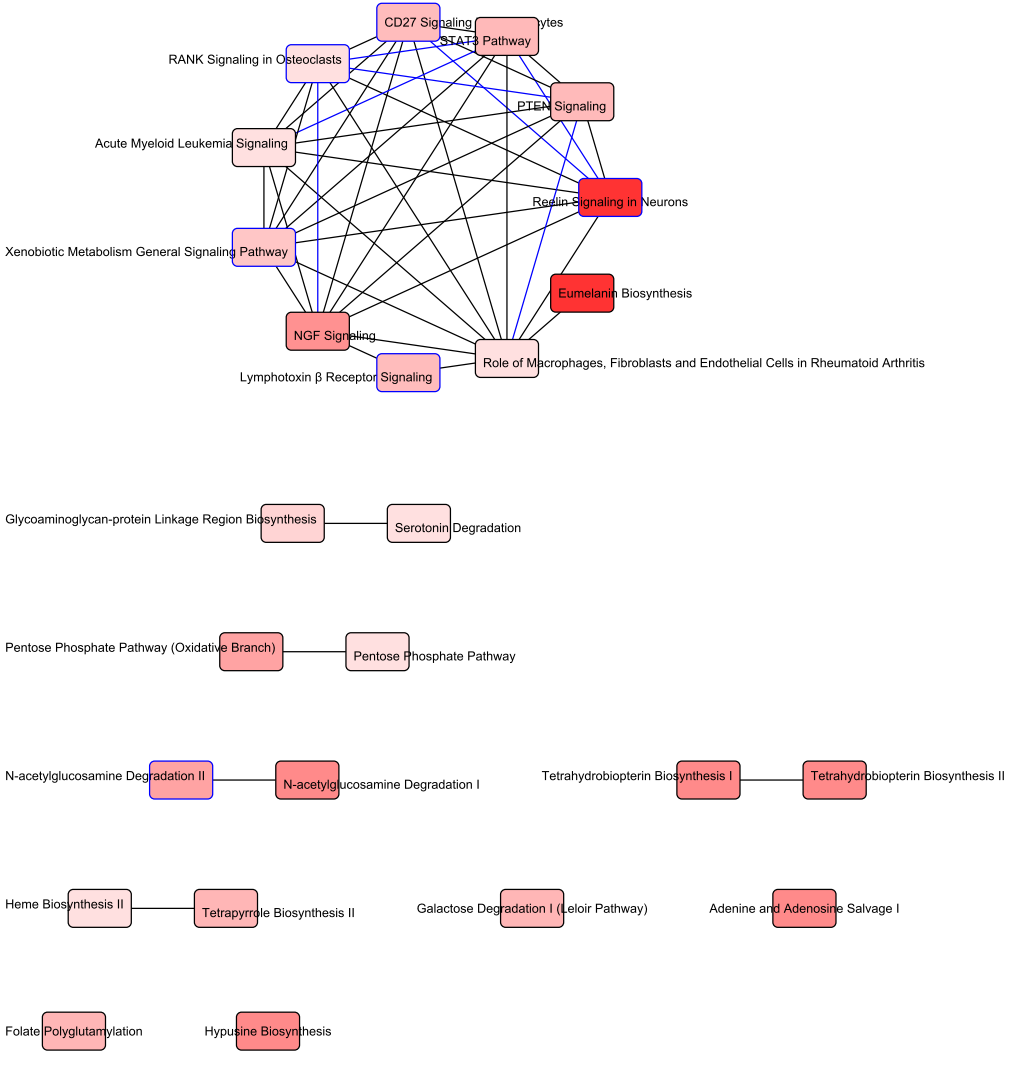

CD27 Signaling
STAT3 Pathway
RANK Signaling in Osteoclasts
PTEN Signaling
Acute Myeloid Leukemia Signaling
Reelin Signaling in Neurons
Xenobiotic Metabolism General Signaling Pathway
Eumelanin Biosynthesis
NGF Signaling
Role of Macrophages, Fibroblasts and Endothelial Cells in Rheumatoid Arthritis
Lymphotoxin β Receptor Signaling
Glycoaminoglycan-protein Linkage Region Biosynthesis
Serotonin Degradation
Pentose Phosphate Pathway (Oxidative Branch)
Pentose Phosphate Pathway
N-acetylglucosamine Degradation II
N-acetylglucosamine Degradation I
Tetrahydrobiopterin Biosynthesis I
Tetrahydrobiopterin Biosynthesis II
Heme Biosynthesis II
Tetrapyrrole Biosynthesis II
Galactose Degradation I (Leloir Pathway)
Adenine and Adenosine Salvage I
Folate Polyglutamylation
Hypusine Biosynthesis

Table 2: Top-ranked sub-networks identified by QIAGEN IPA software for WT-C vs. WT-D. Up-regulated and down-regulated molecules are shown with red and green arrows, respectively.

Molecules in Network	Top Diseases and Functions
1. ↑ ASPM, ↑ BRCA1, ↑CCNB1, ↑ CDCA5, ↑ CDK1, ↑ CENPF, ↑ CEP55, Cop9 Signalosome, ↑ CYTH3, ↑ DLGAP5	Cell Cycle, Cellular Assembly and Organization
2. ↑ DTX3L, ↑ GBP3, ↑ GBP5, ↑ GBP6, ↑ GBP7, ↑ HERC6, ↑ Ifi27l2a/Ifi27l2b, ↑IFI44, ↑ IFIT1B, ↑ IFIT2	Antimicrobial Response, Immunological Disease
3. Abl1/2, ↑ ADAMTSL5, ↑ ARHGAP19, ↓ CALML4, ↑ Clec2d (includes others), ↑ COL12A1, ↑ COL15A1,	Cancer, Connective Tissue Disorders Organ
4. ↓ Akr1c14, ↓ ANGPTL7, ↑ ARHGAP23, ↑ ARHGEF17, ↑ BMPER, ↑ C1QTNF1, ↑ CDC42EP4, ↑ CHST1, ↑ CHST15	Carbohydrate Metabolism, Connective Tissue
5. 14-3-3, ↓ ACAT1, ↑ CYTH4, EGLN, ↑ EPAS1, ↑ ESYT1, CUFLNC, ↓ GUCD1, ↑HK2, ↑ ITGA4	Carbohydrate Metabolism, Cardiovascular System
6. ↑ ACSBG1, ↓ AKIP1, Akt, ↑ANGPTL2, ↑ ASTN2, ↑ ATP1B2, ↑ C1QC, ↑ C9orf116, ↑ CD93, ↑ CMTM7	Carbohydrate Metabolism, Molecular Transport
7. ↑ ALPK1, ↑ C1QTNF7, CD80/CD86, ↑ CLEC12a, ↑ CMPK2, ↑ Cxcl11, ↑ DDX58, ↑ DDX60, ↑EPSTI1	Antimicrobial Response Immunological Disease
8. ↑ ACTC1, ↑ ATAD2, ↑ ATP8B2, ATPase, Calmodulin, Cathepsin, ↑ CKAP2L, ↑ CTSC, ↓ CTSH, ↑ CTSK	Cellular Assembly and Organization
9. ↑ ANXA6, ↓ AS3MT, calpain, ↑ CAPNS, ↑ CAPN6, Cyclin E, ↑DENND2A, ↑ DLG4, ↑ EDN1, Fgfr	Cell Morphology, Cell-To-Cell Signaling
10. ↑ ACTA2, ↑ ACTN1, ↑ ANXA3, ↑ CAV1, ↑ CD44, Creb, ↑CSRP1, cytochrome-c oxidase, ↑ EFHC2, ↓ Gimap9	Cardiovascular System Development
11. ↑ ACAP2, ↑ ARHGAP28, ↑ BIRC3, CD3, ↑ CD300LD, ↑Ear2 (includes others), ↓ EBP, ↑ FIGNL1, ↑ FPR2, ↑ FTCD	Cell-To-Cell Signaling and Interaction
12. ↑ AEN, ↑ ANLN, ↓ APOM, ↑ ARHGAP11A, ↓ ARSG, AURK, ↑ BCHE, ↑ BTG2, ↑ C1orf198, Caspase 3/7	Cellular Assembly and Organization, Cellular Signaling
13. ↑ CDKL5, ↑ CLU, ↑ DLG2, ↑ DLGAP4, ↑ DSCAML1, ↑ FNY, glutathione transferase, Glutathione-S-transferase	Cell-To-Cell Signaling and Interaction
14. ↑ ACSF2, ↓ ACSM5, Adaptor protein 1, adhesion molecule, Aldose Reductase, ↑ ALOX5, ↑ DOCK2, ↓ E2F5, Eif2	Nervous System Development and Function
15. ADRB, Alpha 1 antitrypsin, ↑ANKRD1, ↑ASNS, ↑ ATF3, ↑ C3, ↑ C5AR1, ↓ CA4, CaMKII, ↑ DDIT3	Cell Death and Survival, Organismal Injury
16. ↓ ABHD14A, ↑ ADAR, atypical protein kinase C, ↑ AXL, BCR (complex), ↑ CLEC6A, ↑ CSF1, ↑ EFHD1, ↑ EPB41L2	Cellular Movement, Hematological System Defect
17. ↑ ADAMTSL2, ↑ BMP1, ↑ CASP14, ↑ CCN4, ↑ COL1A2, ↑ COL5A2, ↑ COL6A1, ↑ COL6A2, collagen Collagen Alpha1	Dermatological Diseases and Conditions
18. apyrase, ↑ CD200, ↓ CD320, ↑ CYBRD1, ↑ CYP1B1, ↑ CYP2S1, ↓ CYP4B1, ↑ DAAM, ↑ FAM234B, Fcer1	Developmental Disorder, Molecular Transport
19. ↑ ADAM11, ↑ ADAM22, ↑ ADAMTS1, ↑ ADAMTS12, ↑ADAMTS14, ↑ ADAMTS2, ↑ ADAMTS5, ↑ ADAMTS7, ↑ART	Connective Tissue Disorders, Organismal Injury
20. ALT, ↑ ARHGAP45, ↑ B3GNT7, ↑ CD72, ↑ CLEC9A, ↑ CORO1A, cytokine receptor, ↓ Gm1123, GOT, ↑ HELZ2	Hematological System Development

Figure 6

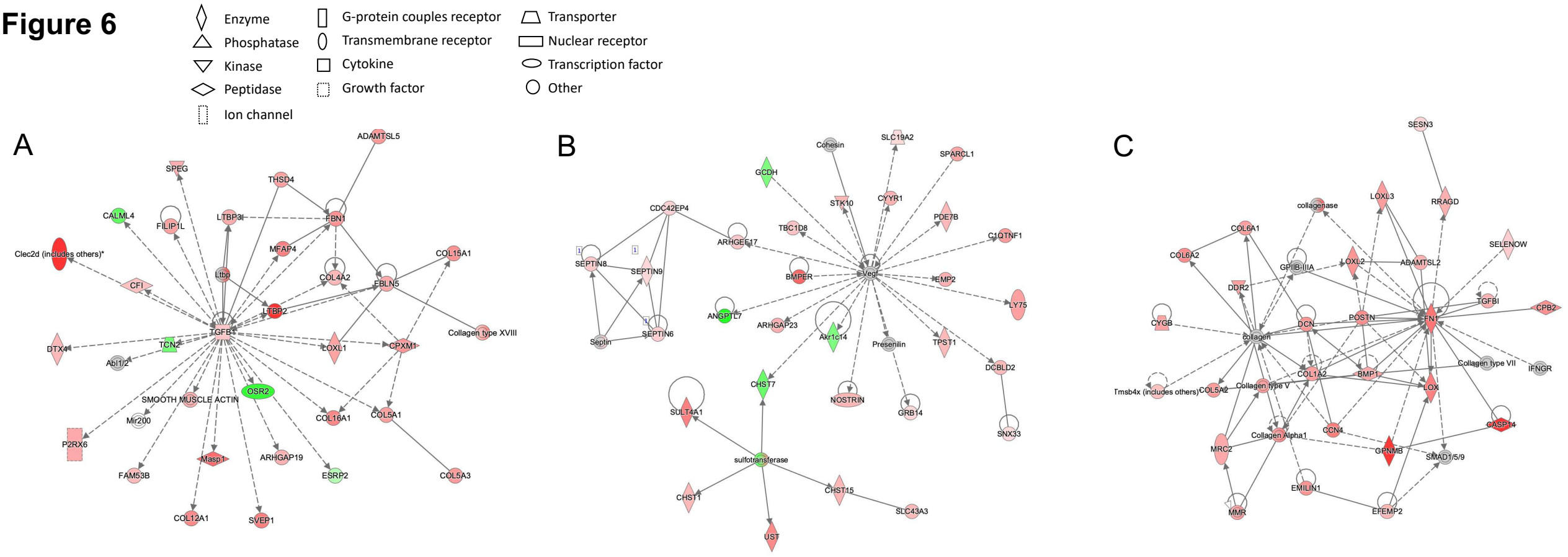
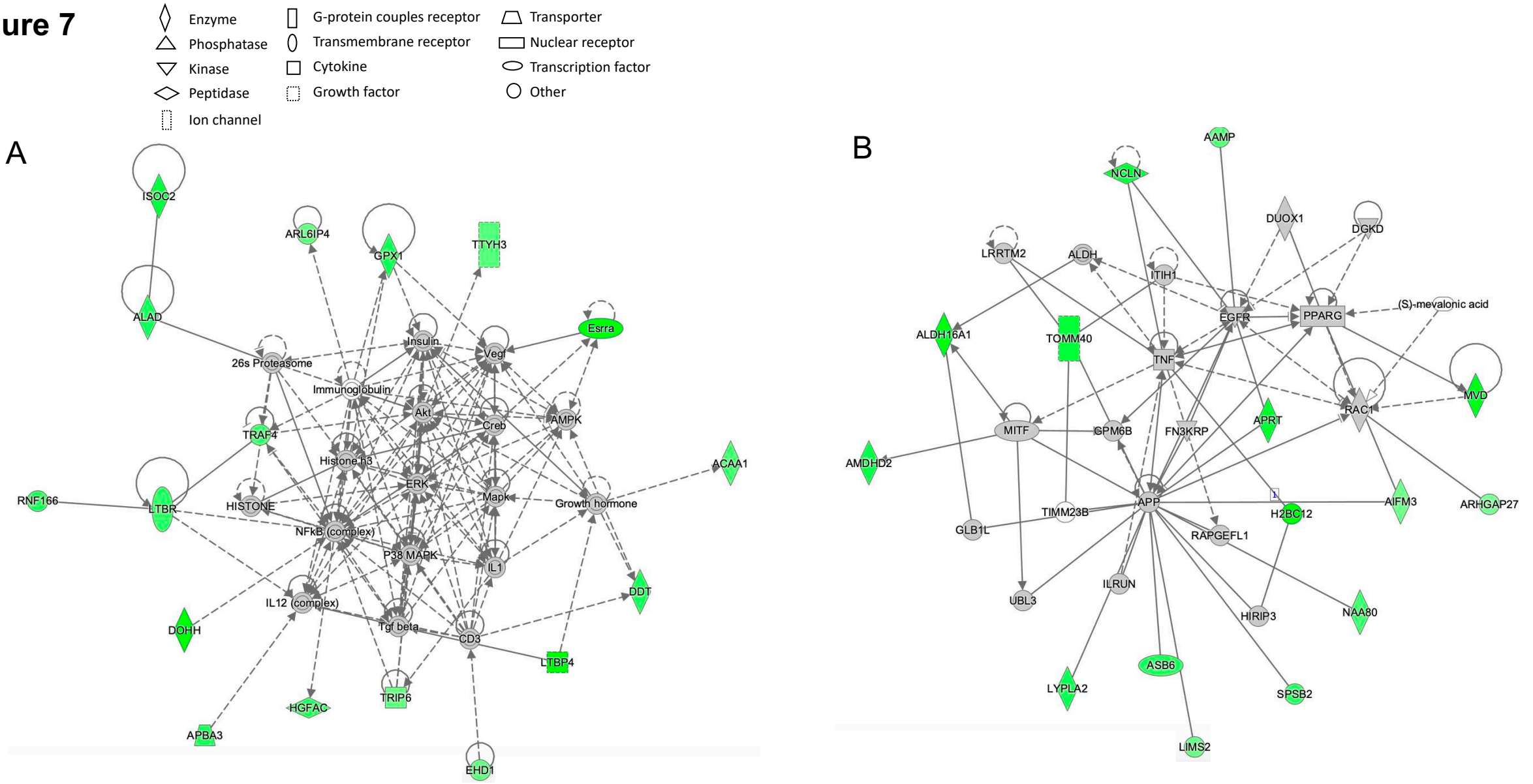

Enzyme
Phosphatase
Kinase
Peptidase
Ion channel
G-protein couples receptor
Transmembrane receptor
Cytokine
Growth factor
Transporter
Nuclear receptor
Transcription factor
Other
A
ADAMTSL5
SPEG
THSD4
CALML4
FILIP1L
LTBP3
FBN1
Clec2d (includes others)*
MFAP4
COL15A1
CFI
Ltbp
COL4A2
FBLN5
LTBP2
TGFB1
Collagen type XVIII
DTX4
TCN2
LOXL1
CPXM1
Abi1/2
SMOOTH MUSCLE ACTIN
OSR2
Mir200
COL16A1
COL5A1
P2RX6
Masp1
ARHGAP19
FAM53B
ESRP2
COL5A3
COL12A1
SVEP1
B
Cohesin
SLC19A2
SPARCL1
GCDH
CDC42EP4
STK10
CYYR1
PDE7B
TBC1D8
C1QTNF1
ARHGEF17
SEPTIN8
SEPTIN9
BMPER
Vegf
EMP2
LY75
ANGPTL7
ARHGAP23
Akr1c14
Presenilin
TPST1
Septin
SEPTIN6
DCBLD2
CHST7
NOSTRIN
GRB14
SULT4A1
SNX33
sulfotransferase
CHST11
CHST15
SLC43A3
UST
C
SESN3
LOXL3
RRAGD
collagenase
COL6A1
COL6A2
GPIIB-IIIA
LOXL2
ADAMTSL2
SELENOW
DDR2
TGFBI
CPB2
CYGB
DCN
POSTN
FN1
collagen
COL1A2
BMP1
Collagen type VII
IFNGR
Tmsb4x (includes others)
COL5A2
Collagen type V
LOX
CASP14
Collagen Alpha1
CCN4
MRC2
GPNMB
SMAD1/5/9
EMILIN1
MMR
EFEMP2

Table 3: Top-ranked sub-networks identified by QIAGEN IPA software for WT-D vs. KO-D. Up-regulated and down-regulated molecules are shown with red and green arrows, respectively.

Molecules in Network	Top Diseases and Functions
1. ↓ABHD14A, ↓ACADS, Ap1, caspase, ↓CBLC, ↓CYBA, EGLN, ↓EGLN2, ↓EMX1, ERK1/2	Developmental Disorder, Hereditary Disorder
2. 26s Proteasome, ↓ACAA1, Akt, ↓ALAD, AMPK, ↓APBA3, ↓ARL6IP4, CD3, Creb, ↓DDT	Hematological System Development and Function
3. (S)-mevalonic acid, ↓AAMP, ↓AIFM3, ALDH, ↓ALDH16A1, ↓AMDHD2, ↓APP, ↓APRT, ↓ARHGAP27, ↓ASB6	Cell morphology, cellular movement, Hematology
4. Aldehyde dehydrogenase (NAD), ↓Aldh3b2, ↓ARHGEF16, ↓B3GAT3, ↓CCDC124, ↑CTNNB1, dihydroxyacetone	Cancer, Organismal Injury and Abnormalities
5. ↓ABHD16A, ↓ ADPRS, ↓Bola2, ↑CNN4, ↓CDK2AP2, ↑CHD2, ↑CHDS, ↑EED, ↑FRAT1, ↓GALK1	Cancer, Neurological Disease
6. ↓ABHD17A, ↓ACKR3, ADCY, adenosine triphosphate, ↑ADGRG3, ↑ APLNR, beta-estrodiol, DRD3, ↓FFAR3, Folate	Cell Signaling, Cell-To-Cell Signaling
7. ↓AKRA7A2, AKR7A3, AKR7L, ↓ASPDH, ↓CANT1, ↓CC2D1B, ↓CCDC124, COMMD5, ↓COMMD9, ↑DENND2C	Cardiovascular Disease, Drug Metabolism

Figure 7


Enzyme
Phosphatase
Kinase
Peptidase
Ion channel
G-protein couples receptor
Transmembrane receptor
Cytokine
Growth factor
Transporter
Nuclear receptor
Transcription factor
Other
A
ISOC2
ALAD
ARL6IP4
GPX1
TTYH3
Esrra
Insulin
Vegf
26s Proteasome
Immunoglobulin
Akt
Creb
AMPK
TRAF4
Histone h3
ERK
Mapk
Growth hormone
ACAA1
RNF166
LTBR
HISTONE
NFkB (complex)
P38 MAPK
IL1
DDT
IL12 (complex)
DOHH
Tgf beta
CD3
LTBP4
TRIP6
HGFAC
APBA3
EHD1
B
AAMP
NCLN
DUOX1
DGKD
LRRTM2
ALDH
ITIH1
(S)-mevalonic acid
EGFR
PPARG
ALDH16A1
TOMM40
TNF
MVD
RAC1
APRT
MITF
GPM6B
FN3KRP
AMDHD2
APP
AIFM3
ARHGAP27
TIMM23B
H2BC12
GLB1L
RAPGEFL1
ILRUN
UBL3
HIRIP3
NAA80
ASB6
LYPLA2
SPSB2
LIMS2

networks of interest. Here we highlighted two networks namely ERK and TNF. These molecules were unchanged, possibly due to the down-regulation of regulating molecules.

Discussion

At the transcription and post transcription levels, a symphony of regulatory molecules including transcription factors, transcription co-activators, DNA and histone methylators, non-coding RNAs and others have been shown to play critical roles in altered protein production [27]. Such epigenetic mechanisms play a major role in all chronic disease processes including DN [7,8].

Long non-coding RNAs with a size of >200 bp play a major role in gene regulation [9]. Long non-coding RNAs, in general, acts through multiple mechanisms and may on a nearby *[cis*] or distant genes [*trans*] regulating their transcription [9,10]. We took a novel approach using RNA sequencing and bioinformatics analysis to answer the role of a specific lncRNA, *ANRIL* in DN.

We have previously demonstrated a pathogenetic role of *ANRIL* on the pathogenesis of several chronic diabetic complications including DN [13]. The major aim of this study was to delineate the mechanisms of such protection.

In this study, using RNA sequencing and bioinformatic analyses approaches, we demonstrated that the vast majority of the altered renal transcripts, in the context of DN, are regulated by *ANRIL* as they were mostly normalized in the diabetic ANRIL KO mice. In the kidneys of wild type diabetic animals, as a result of diabetic dysmetabolism, RNA

transcripts involved in a large number of pathways were changed. The major pathways, as observed in both KEGG and REACTOME analyses, include AGE-RAGE signaling, PI3K-AKT, metabolic pathways, TNFα and NF KB related molecules were changed. These pathways are known to mediate renal damage in DN. However, a large number of additional molecules, yet to be characterized in DN were also changed. In the kidneys of the diabetic KO animals, even in the presence of hyperglycemia, all important pathways were corrected in association with renal functional changes in DN. The exact significance of the remaining molecules, which were not corrected, are not clear. It is of further interest to note that as the animals remained hyperglycemic, the metabolic pathway related genes were not normalized in ANRIL KO mice with diabetes (KOD).

As noted above we have previously reported that the expression of ANRIL was upregulated in DN and in other chronic diabetic complications [12,13]. We have also shown that nullifying ANRIL production prevents several biochemical, functional and structural changes in DN [13]. As planned, the current study delineates molecular mechanisms of ANRIL mediated pathogenesis in DN.

Although the role of ANRIL has been previously explored in the context of cardiovascular disease, very few studies have been conducted with respect to ANRIL in DN. We have demonstrated its role in regulating structural and function changes in DN. It has recently been demonstrated that ANRIL promotes pyroptosis and kidney injury in DN [28].

Based on our data it is possible that ANRIL may be considered as a drug target in DN and possibly other chronic diabetic complications. Based on these findings it is even tempting to speculate that ANRIL may potentially represent a one-stop shop for DN therapy. Such therapy may exploit an RNA based approach or a small molecule inhibitor

targeting ANRIL. As no significant alterations of transcripts were seen in the non-diabetic ANRIL KO animals, such RNA targeting therapy conceptually may lead to limited side effects. However, such notion needs to be initially confirmed using long term large scale preclinical studies.

In summary, we have demonstrated, using an RNA sequencing and bioinformatics approach, that alterations of a large number of molecules associated with DN are regulated through lncRNA ANRIL. These finding suggest that ANRIL may potentially be a drug target for DN and possibly of other chronic diabetic complications:

Competing interests

The author(s) declare no competing interests.

Author contributions

Experimental conception and design: P.S., S.C. Performed the experiments: B.F., S.B., H., L. Reagents/ materials/analysis tools contribution: P.S., M.L., H. L., Z. S. Writing of manuscript: S.C., P.S. Manuscript review: all.

Acknowledgement:

Supported by from the Canadian Institutes of Health Research (funding reference number: 169650) (SC), Schulich School of Medicine and Dentistry, Western University, Lawson Internal Research Fund (PS) and Jiangsu Province 100 talent International collaborative research program (BX2019100) (SC and ZS). PS is a recipient of new investigator award from Ontario Institute for Cancer Research (OICR).

References:

1. Gallagher, H. and Suckling, R.J. Diabetic nephropathy: where are we on the journey from pathophysiology to treatment? *Diabetes, Obes and Metab* 18: 641-647 (2016).

2. Radica, Z., Alicic, T. R., and Katherine R. T. Diabetic Kidney Disease Challenges, Progress, and Possibilities. CJASN, 12, 2032-2045 (2017)

3. Ghaderian, S., B., Hayati., F., Shayanpour, S.,, Mousavi, S., S., B. Diabetes and end-stage renal disease; a review article on new concepts. *J Renal Inj Prev* 4: 28-33, (2015).

4. Komers, R., et.al. Transcriptome-Based Analysis of Kidney Gene Expression Changes Associated with Diabetes in OVE26 Mice, in the Presence and Absence of Losartan Treatment https://doi.org/10.1371/journal.pone.0096987

5. Benjamin, L., E. Glucose, VEGF-A, and Diabetic Complications. *Am J Pathol*. 158: 1181-1184, (2001).

6. Eddy AA. Molecular basis of renal fibrosis*. Pediatr Nephrol 15: 290-301, (2000).*

7. Portela, A., and Esteller, M. Epigenetic modifications and human disease. *Nature Biotechnol 28: 1057-1068 (2010).*

8. Thomas, M.,C. Epigenetic mechanisms in diabetic kidney disease. *Curr Diab Rep 16:1-8, (2016).*

9. Wapinsk., I.,O. and Chang HY. Long noncoding RNAs and human disease. *Trends Cell Biol 21: 354-361 (2011).*

10. Aguilo, F., Di Cecilia, S., and Walsh, M.,J. Long Non-coding RNA ANRIL and Polycomb in Human Cancers and Cardiovascular Disease. *Curr Top Microbiol Immunol 394: 29-39, (2016).*

11. Congrains A, Kamide K, Ohishi M, Rakugi H. ANRIL: molecular mechanisms and implications in human health. Int J Mol Sci 14: 1278-1292, 2013.

12. Thomas, A., A., Feng, B., and Chakrabarti, S. ANRIL: A Regulator of VEGF in Diabetic Retinopathy*. Invest Ophthalmol Vis Sci 58: 470-480, (2017)*

13. Thomas., A., A., Feng, B.,, and Chakrabarti*,* S. ANRIL regulates production of extracellular matrix proteins and vasoactive factors in diabetic complications *Am J Physiol Endocrinol Metab 314: E191–E200, (2018).*

14. Wan, G., et.al. Long non-coding RNA ANRIL (CDKN2B-AS) is induced by the ATM-E2F1 signaling pathway. *Cell Signal. 25:1086-95 (2013).*

15. Visel, Aet.al. Targeted deletion of the 9p21 non-coding coronary artery disease risk interval in mice. Nature 464: 409-412, 2010.

16. Wang, C., George, B., Chen, S., Feng, B., Li, X., Chakrabarti, S.. Genotoxic stress and activation of novel DNA repair enzymes in the endothelial cells and in the retinas and kidneys in diabetes. *Diabetes Metab Res Rev.* doi: 10.1002/dmrr.2279. (2012)

17. Feng, B., et.al. miR-200b Mediates Endothelial-to-Mesenchymal Transition in Diabetic Cardiomyopathy. *Diabetes*; 65:1-11 (2016).

18. Feng, B., Cao, Y., Chen, S., Ruiz, M., and Chakrabarti, S. miRNA-1 regulates endothelin-1 in diabetes. *Life Sci* 98: 18-23, (2014).

19. Dobin, A.,. et.al.. "STAR: ultrafast universal RNA-seq aligner." *Bioinformatics 29(1): 15-21 (2013).*

20. Lawrence, M., et,al . "Software for computing and annotating genomic ranges." *PLoS Comput Biol 9(8): e1003118 (2013).*

21. Love, M., I., Huber, W., and S. Anders, S. "Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2." *Genome Biol 15: 550 (2014).*

22. Robinson, M., D., McCarthy, D., J., and Smyth G., K. "edgeR: a Bioconductor package for differential expression analysis of digital gene expression data." *Bioinformatics 26: 139-140 (2010).*

23. Kanehisa, M. and Goto, S.; KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res. 28, 27-30 (2000).

24. Kanehisa, M; Toward understanding the origin and evolution of cellular organisms. Protein Sci. 28, 1947-1951 (2019)

25. Kanehisa, M., Furumichi, M., Sato, Y., Ishiguro-Watanabe, M., and Tanabe, M.; KEGG: integrating viruses and cellular organisms. Nucleic Acids Res. 49, D545-D551 (2021).

26. Jassal B, Matthews L, Viteri G, Gong C, Lorente P, Fabregat A, Sidiropoulos K, Cook J, Gillespie M, Haw R, Loney F, May B, Milacic M, Rothfels K, Sevilla C, Shamovsky V, Shorser S, Varusai T, Weiser J, Wu G, Stein L, Hermjakob H, D'Eustachio P. The reactome pathway knowledgebase. *Nucleic Acids Res.* 2020 Jan 8;48(D1):D498-D503. doi: 10.1093/nar/gkz1031.

27. Mitsuo Kato, M., Natarajan, R. Epigenetics and epigenomics in diabetic kidney disease and metabolic memory. *Nat Rev Nephrol. 15: 327–345 (2019)*

28. Su, J., W., Zhao, M. LncRNA-antisense non-coding RNA in the INK4 locus promotes pyroptosis via miR-497/thioredoxin-interacting protein axis in diabetic nephropathy. *Life Sciences DOI: 10.1016/j.lfs.2020.118728).*

Figure texts:

Figure 1: Diabetes-induced A) elevated blood glucose levels, B) reduced body weight and C) increased urinary albumin/creatinine Ratio were prevented in the diabetic animals lacking ANRIL (KO-D). (WT-C= wild type controls, WT-D wild-type diabetic, KO-C=ANRIL knockout controls, *P = 0.05 or less vs. WT-C, ** P = 0.05 or less vs. KO-C).

Figure 2: Number of differentially expressed genes identified by DESeq2 (light grey) and EdgeR (red) [WT= wild-type, D= diabetes, C= age matched non-diabetic controls, KO = Knockout). Data are presented A) at a p value threshold of 0.05 and B) at a p value threshold of 0.01. Please note that minimal number of alterations occurred at the basal level (WT-C vs KO-C). Poorly controlled diabetes caused alterations of a large number of transcripts in the wild type (WT) animals. ANRIL knockout (KO) prevented a large number of such alterations.

Figure 3: Top-ranked pathways and number of significant genes from each pathway. (A) Significant KEGG pathways enriched for genes differentially expressed (adjusted p < 0.01) between wild-type controls and wild-type diabetes mice (WT-C vs. WT-D), and (B) the number of down-regulated (green), up-regulated (red) and non-significant (white) genes mapped to each pathway. (C) KEGG pathway analysis showed that metabolic pathway is enriched for genes differentially expressed (adjusted p < 0.01) between wild-type diabetes and knockout mice (WT-D vs. KO-D), and (D) *none* of the genes mapped to KEGG metabolic pathway is up-regulated. (E) List of top Reactome pathways for (WT-C vs. WT-D), and (F) the number of genes mapped to each pathway. Down-regulated and up-regulated genes are shown in green and red, respectively.
Both KEGG (A) and Reactome (E) pathway analyses showed alterations (p<0.01) of transcripts related to multiple biological pathways in the kidneys in diabetes. The majority of these, except for the transcripts related to metabolic pathways in KEGG analyses (C), were corrected in the ANRIL KO mice. Reactome analyses didn't detect any changes between WT-D vs. KO-D. (C= non-diabetic control, D= poorly controlled diabetic, WT = wild type, KO= ANRIL knockout, *= some pathways known to alter in diabetic kidney diseases, Horizontal line = FDR of 0.01, the analyses at the level of P<0.05 have been depicted in the supplementary Figure 1. The detailed listings of these transcripts are in the supplementary tables (Suppl. Table 1, Suppl. Table 2 and Suppl. Table 3).

Figure 4: Cluster analyses showing when comparing (A) Wild-type diabetic with wild-type non-diabetic controls (WT-D vs. WT-C), the differentially expressed transcripts were organized in at least 4 major clusters (with > 25 genes within each cluster). However, when (B) wild-type diabatic (WT-D) animals were compared to ANRIL knockout diabetes (KO-D), very few transcripts remain differentially expressed, and the cluster sizes were smaller. Suppl. Table 4 provides detailed listings of the constituents of these cluster.

Figure 5: Overlaps of enriched IPA canonical pathways. Interactions of enriched canonical pathways are shown for (A) wild-type controls vs. diabetes (WT-C vs. WT-D) and (B) wild-type diabetes vs. knockout diabetes (WT-D vs. KO-D). Each node corresponds to one canonical pathway, and the lines connecting pathways indicate

interactions. Nodes (i.e. pathways) are colored based on their significance levels; such that darker nodes represent more significant pathways. Threshold of significance was defined as FDR < 0.01. Please note that the WT-C vs. WT-D pathways are more significant and they are tightly connected, while WT-D vs. KO-D are less significant (lighter colors) and less connected.

Figure 6: Gene interaction network maps. In the network, each gene or gene product is presented by a node, and the biological relationship between two nodes is presented by an edge between the two nodes. Color of the nodes indicate if they are down-regulated (green) or up-regulated (red). The darker colors show more significant genes and the lighter colors present less significant ones. Uncolored nodes represent genes that are not differentially expressed at a significant level in our dataset. However, they are presented in the network based on the evidence in the Ingenuity Pathways Knowledge Base, supporting their relationship to other genes in the network. The node shapes indicate their functions. The networks outlined here represent 3, 4 and 17 of Table 2, corresponding to A) TGFβ1, B) VEGF and C) collagen molecules. We selected these networks as these are well established molecules of pathogenetic significance in DN.

Figure 7: Gene interaction network maps. In the network, each gene or gene product is presented by a node, and the biological relationship between two nodes is presented by an edge between the two nodes. Color of the nodes indicate if they are down-regulated (green) or up-regulated (red). Uncolored nodes represent genes that are not differentially expressed at a significant level in our dataset. However, they are presented in the network based on the evidence in the Ingenuity Pathways Knowledge Base, supporting their relationship to other genes in the network. The node shapes indicate their functions. Here we highlighted two networks namely A) ERK and B) TNF. These molecules were unchanged, possibly due to the down-regulation of regulating molecules.

Figures

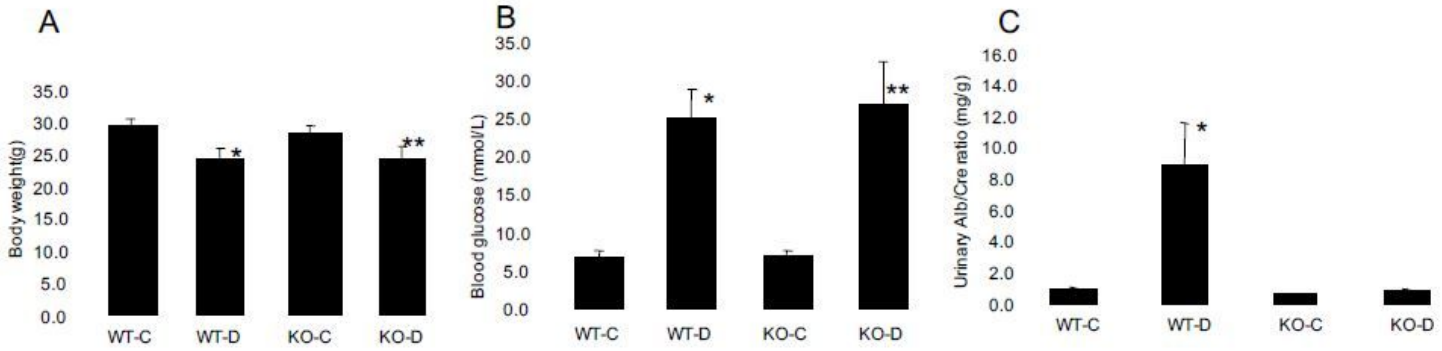


Figure 1

Diabetes-induced A) elevated blood glucose levels, B) reduced body weight and C) increased urinary albumin/creatinine Ratio were prevented in the diabetic animals lacking ANRIL (KO-D). (WT-C= wild type controls, WT-D wild-type diabetic, KO-C=ANRIL knockout controls, *P = 0.05 or less vs. WT-C, ** P = 0.05 or less vs. KO-C).

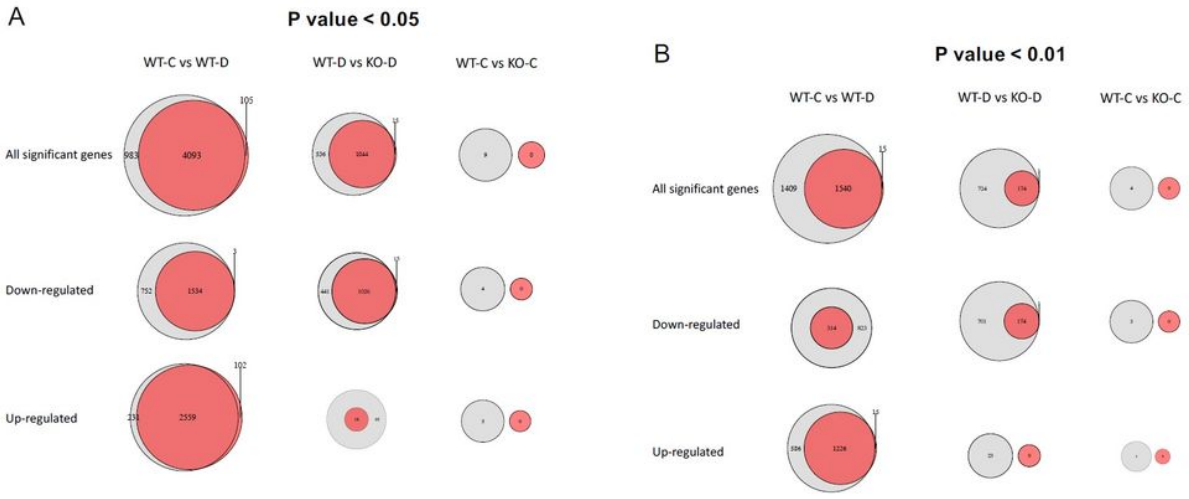


Figure 2

Number of differentially expressed genes identified by DESeq2 (light grey) and EdgeR (red) [WT= wild-type, D= diabetes, C= age matched non-diabetic controls, KO = Knockout). Data are presented A) at a p value threshold of 0.05 and B) at a p value threshold of 0.01. Please note that minimal number of alterations occurred at the basal level (WT-C vs KO-C). Poorly controlled diabetes caused alterations of a

large number of transcripts in the wild type (WT) animals. ANRIL knockout (KO) prevented a large number of such alterations.

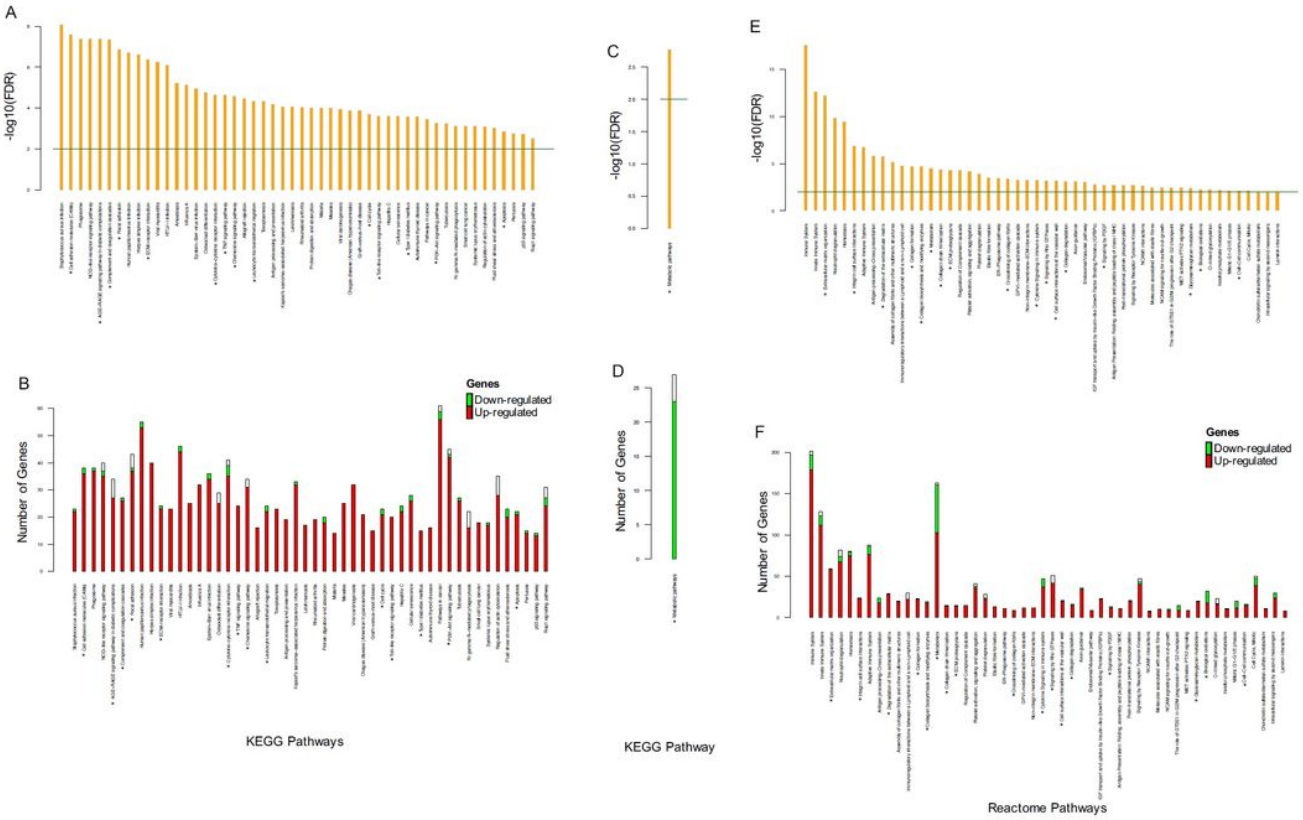


Figure 3

Top-ranked pathways and number of significant genes from each pathway. (A) Significant KEGG pathways enriched for genes differentially expressed (adjusted p < 0.01) between wild-type controls and wild-type diabetes mice (WT-C vs. WT-D), and (B) the number of down-regulated (green), up-regulated (red) and non-significant (white) genes mapped to each pathway. (C) KEGG pathway analysis showed that metabolic pathway is enriched for genes differentially expressed (adjusted p < 0.01) between wildtype diabetes and knockout mice (WT-D vs. KO-D), and (D) none of the genes mapped to KEGG metabolic pathway is up-regulated. (E) List of top Reactome pathways for (WTC vs. WT-D), and (F) the number of genes mapped to each pathway. Down-regulated and up-regulated genes are shown in green and red, respectively. Both KEGG (A) and Reactome (E) pathway analyses showed alterations (p<0.01) of transcripts related to multiple biological pathways in the kidneys in diabetes. The majority of these, except for the transcripts related to metabolic pathways in KEGG analyses (C), were corrected in the ANRIL KO mice. Reactome analyses didn't detect any changes between WT-D vs. KO-D. (C= non-diabetic control, D= poorly controlled diabetic, WT = wild type, KO= ANRIL knockout, *= some pathways known to alter in diabetic kidney diseases, Horizontal line = FDR of 0.01, the analyses at the level of P<0.05 have

been depicted in the supplementary Figure 1. The detailed listings of these transcripts are in the supplementary tables (Suppl. Table 1, Suppl. Table 2 and Suppl. Table 3).

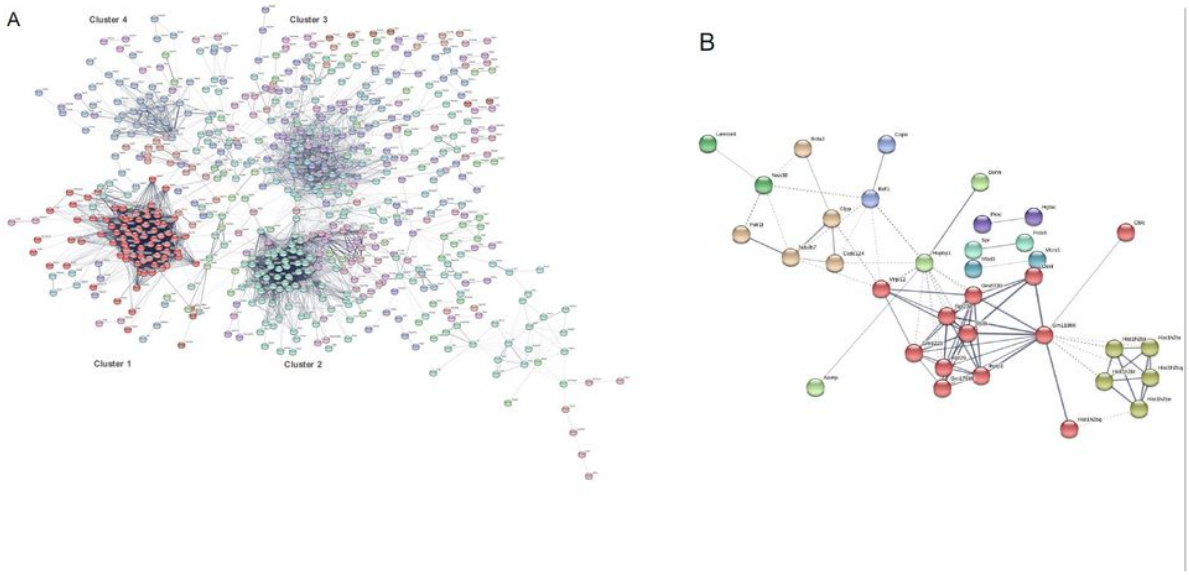


Figure 4

Cluster analyses showing when comparing (A) Wild-type diabetic with wild-type non-diabetic controls (WT-D vs. WT-C), the differentially expressed transcripts were organized in at least 4 major clusters (with > 25 genes within each cluster). However, when (B) wild-type diabatic (WT-D) animals were compared to ANRIL knockout diabetes (KO-D), very few transcripts remain differentially expressed, and the cluster sizes were smaller. Suppl. Table 4 provides detailed listings of the constituents of these cluster.

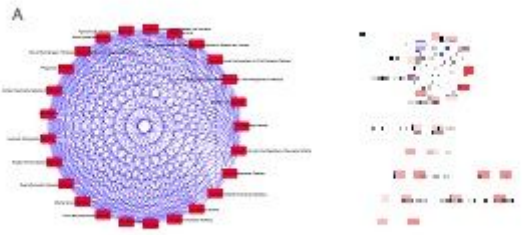


Figure 5

Overlaps of enriched IPA canonical pathways. Interactions of enriched canonical pathways are shown for (A) wild-type controls vs. diabetes (WT-C vs. WT-D) and (B) wild-type diabetes vs. knockout diabetes (WT-D vs. KO-D). Each node corresponds to one canonical pathway, and the lines connecting pathways indicate 23 interactions. Nodes (i.e. pathways) are colored based on their significance levels; such that darker nodes represent more significant pathways. Threshold of significance was defined as FDR < 0.01. Please note that the WT-C vs. WT-D pathways are more significant and they are tightly connected, while WT-D vs. KO-D are less significant (lighter colors) and less connected.

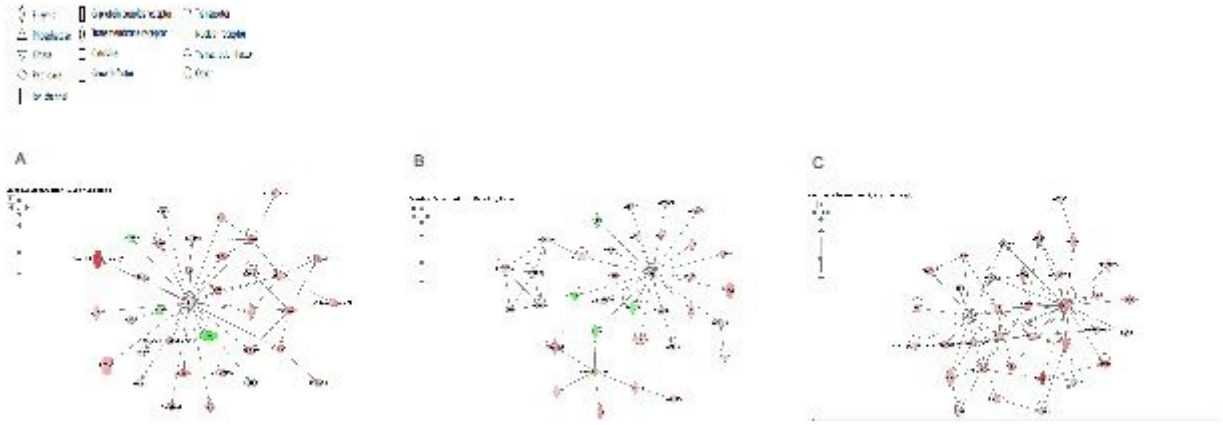


Figure 6

Gene interaction network maps. In the network, each gene or gene product is presented by a node, and the biological relationship between two nodes is presented by an edge between the two nodes. Color of the nodes indicate if they are down-regulated (green) or up-regulated (red). The darker colors show more significant genes and the lighter colors present less significant ones. Uncolored nodes represent genes that are not differentially expressed at a significant level in our dataset. However, they are presented in the network based on the evidence in the Ingenuity Pathways Knowledge Base, supporting their relationship to other genes in the network. The node shapes indicate their functions. The networks outlined here represent 3, 4 and 17 of Table 2, corresponding to A) TGFβ1, B) VEGF and C) collagen molecules. We selected these networks as these are well established molecules of pathogenetic significance in DN.

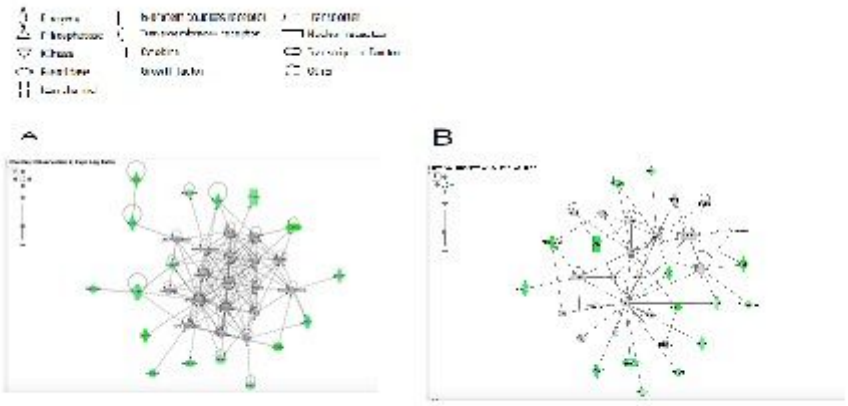


Figure 7

Gene interaction network maps. In the network, each gene or gene product is presented by a node, and the biological relationship between two nodes is presented by an edge between the two nodes. Color of the nodes indicate if they are down-regulated (green) or up-regulated (red). Uncolored nodes represent genes that are not differentially expressed at a significant level in our dataset. However, they are presented in the network based on the evidence in the Ingenuity Pathways Knowledge Base, supporting their relationship to other genes in the network. The node shapes indicate their functions. Here we highlighted two networks namely A) ERK and B) TNF. These molecules were unchanged, possibly due to the down-regulation of regulating molecules.

Supplementary Files

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

- Supplfigtext.pdf
- SupplFigure1.pdf
- SupplFigure2.pdf
- SupplTable1.pdf
- SupplTable2.pdf
- SupplTable3.pdf
- SupplTable4.xls