

Published in final edited form as:

Biochim Biophys Acta. 2012 April ; 1822(4): 589–599. doi:10.1016/j.bbadis.2012.01.008.

Next-Generation Sequencing Identifies TGF- β 1-Associated Gene Expression Profiles in Renal Epithelial Cells Reiterated in Human Diabetic Nephropathy

Eoin P. Brennan^{1,a}, Melissa J. Morine^{2,a}, David W. Walsh^{1,a}, Sarah A. Roxburgh¹, Maja T. Lindenmeyer³, Derek P. Brazil⁴, Peadar Ó. Gaora⁵, Helen M. Roche², Denise M. Sadlier¹, Clemens D. Cohen³, The GENIE Consortium⁶, Catherine Godson^{1,b}, and Finian Martin^{1,b,*}

¹UCD Diabetes Research Centre, UCD Conway Institute of Biomolecular & Biomedical Research, University College Dublin, Dublin 4, Ireland ²Nutrigenomics Research Group, UCD Conway Institute of Biomolecular & Biomedical Research, University College Dublin, Dublin 4, Ireland ³Division of Nephrology, University Hospital Zurich, Zurich, Switzerland ⁴Centre for Vision and Vascular Science, Queen's University Belfast, Belfast, UK ⁵School of Medicine and Medical Sciences, UCD Conway Institute of Biomolecular & Biomedical Research, University College Dublin, Dublin 4, Ireland

Abstract

Transforming growth factor-beta (TGF- β 1) is implicated in the onset and progression of renal fibrosis and diabetic nephropathy (DN), leading to a loss of epithelial characteristics of tubular cells. The transcriptional profile of renal tubular epithelial cells stimulated with TGF- β 1 was assessed using RNA-Seq, with 2027 differentially expressed genes identified. Promoter analysis of transcription factor binding sites in the TGF- β 1 responsive gene set predicted activation of multiple transcriptional networks, including NF κ B. Comparison of RNA-Seq with microarray data from identical experimental conditions identified low abundance transcripts exclusive to RNA-Seq data. We compared these findings to human disease by analyzing transcriptomic data from renal biopsies of patients with DN versus control groups, identifying a shared subset of 179 regulated genes. ARK5, encoding an AMP-related kinase, and TGFBI - encoding transforming growth factor, beta-induced protein were induced by TGF- β 1 and also upregulated in human DN. Suppression of ARK5 attenuated fibrotic responses of renal epithelia to TGF- β 1 exposure; and silencing of TGFBI induced expression of the epithelial cell marker – E-cadherin. We identified low abundance transcripts in sequence data and validated expression levels of several transcripts

© 2011 Elsevier B.V. All rights reserved.

*Corresponding author at: Prof. Finian Martin, UCD Diabetes Research Centre, UCD Conway Institute of Biomolecular & Biomedical Research, University College Dublin, Dublin 4, Ireland. finian.martin@ucd.ie.

^aE.P.B., M.J.M. and D.W.W. contributed equally to this work

^bCG and FM share senior authorship of this work

⁶A full list of members is provided in the Supplementary Details

Disclosure

The authors declare no conflict of interest.

Author contributions: F.M., C.G., D.P.B. and D.M.S. designed research; E.P.B. and S.A.R. performed research; D.W.W., M.J.M. and P.ÓG analyzed data; and F.M., C.G., E.P.B., M.J.M., D.W.W., P.ÓG. and H.M.R. wrote the paper. Data deposition: Transcriptomic data described in this paper have been deposited in the Gene Expression Omnibus (GEO) database, www.ncbi.nlm.nih.gov/geo (HK-2 array: accession no. GSE23338; RNA-Seq: accession no. GSE29660).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

(ANKRD56, ENTPD8) in tubular enriched kidney biopsies of DN patients versus living donors. In conclusion, we have defined a TGF- β 1-driven pro-fibrotic signal in renal epithelial cells that is also evident in the DN renal transcriptome.

Keywords

Diabetic nephropathy; renal tubule; epithelial-mesenchymal transition; transforming growth-factor-beta1

1. Introduction

Diabetic nephropathy (DN) is the leading cause of end stage renal disease worldwide, affecting approximately 25–30% of those with diabetes [1]. With the escalating prevalence of diabetes there is a significant rise in the incidence of DN and other diabetic microvascular complications. DN reflects the convergence of metabolic and hemodynamic insults, renal podocyte loss, albuminuria, mesangial cell hypertrophy, glomerulosclerosis and tubulointerstitial fibrosis (TIF) [2]. There is an urgent need to better understand the molecular basis underlying the initiation and progression of DN with a view to defining new biomarkers and targets for therapeutic intervention. TIF is proposed to develop through proliferation and activation of resident fibroblasts, recruitment of circulating fibrocytes and by transformation of epithelial and other resident cells to a mesenchymal phenotype (EMT - epithelial-mesenchymal transformation) [3].

Transforming growth factor- β 1 (TGF- β 1) is implicated in the development of diabetic glomerulosclerosis and there is a more recent appreciation of its role as a key driver of TIF [2, 4]. TGF- β 1 is a member of the TGF- β superfamily of proteins, consisting of three ligands (TGF- β 1, 2 and 3) together with the activins/inhibins and the bone morphogenic proteins. This family is involved in a variety of cell processes including the control of cell growth, proliferation, differentiation and apoptosis [5]. TGF- β 1 signals through serine-threonine kinase type I and II transmembrane receptors, initiating receptor I phosphorylation and signal transduction via activation of canonical (SMAD) and non-canonical (p38MAPK, ERK, JNK, PI3K/Akt) pathways [6, 7]. TGF- β 1-induced TIF can be modeled *in vitro* in cultured renal epithelial cells which de-differentiate into a more mesenchymal-like phenotype in response to the cytokine. Such responses are characterized by a switch in predominant cadherins from E-Cadherin (epithelial) to N-cadherin (mesenchymal), and increased vimentin, α -smooth muscle actin, CTGF and the Notch ligand Jagged [8].

The current study sought to gain greater insight in relation to dysregulated TGF- β 1 activation in DN using RNA Seq. Thus differential transcriptional responses to TGF- β 1 stimulation of human renal epithelial (HK-2) cells was investigated by RNA-Seq; with comparative sensitivity analysis to DNA microarrays. The accuracy and sensitivity of DNA microarrays can be limited by probe coverage restrictions and relative insensitivity to low abundance transcripts [9]. Sequence-based approaches overcome many of these limitations, and here we demonstrate that RNA-Seq identifies a considerably larger number of TGF- β 1-regulated genes.

Using these *in vitro* datasets, we compared TGF- β 1-regulated genes identified by RNA-Seq with available expression data from micro-dissected human kidney biopsies, at various stages of DN versus healthy controls [10], and defined a common subset of 179 TGF- β 1-related genes, comprising both established and potentially novel regulators of renal fibrosis. By silencing the expression of several of these genes, we have demonstrated their functional

significance to the TGF- β 1 fibrotic signal in renal epithelial cells. Finally, we tested whether the large set of TGF- β 1-regulated low-abundance transcripts identified in the cell model sequence dataset could inform the selection of novel transcripts for detection in human DN biopsy samples. Using this strategy, we have identified several genes not previously available on our human DN microarray platforms, and successfully validated these responses in human DN biopsy samples versus living donors.

2. Materials and methods

2.1. Cell culture

Human kidney epithelial cells (HK-2) were purchased from the European Tissue Culture Collection (ETCC) HK-2 cells were cultured at 37 °C in a humidified atmosphere of 95% air/5% CO₂, and maintained in DMEM-F12 (Sigma) supplemented with 2 mM l-glutamine, 100 U/ml penicillin, 100 μ g/ml streptomycin, 10 ng/ml EGF, 36 ng/ml hydrocortisone, 3 pg/ml triiodothyronine and 5 μ g/ml insulin– 5 μ g/ml transferrin– 5ng/ml selenium (ITS) solution (Sigma-Aldrich). Cells were subcultured using trypsin–EDTA at a ratio of 1:5. HK-2 cells were plated in triplicate for stimulation with TGF- β 1 (5 ng/ml PromoCell) or vehicle for 48 hr. Three independent experiments were performed. Triplicate samples for each time point being pooled on each occasion, giving a total of 6 samples for RNA-Seq analysis. Total RNA extraction was performed using an RNeasy RNA extraction kit according to the manufacturer's protocol (Qiagen). RNA quality was assessed using a Bioanalyzer 2100 (Agilent).

2.2. Western blot and qRT-PCR

Lysates were harvested in RIPA lysis buffer containing 50 mM Tris-HCl, pH 7.4, 1% Nonidet P-40, 0.25% sodium deoxycholate, 150 mM NaCl, and 1 mM EDTA, supplemented with 1 mM phenylmethylsulfonyl fluoride, 1 mM sodium orthovanadate, 1mM sodium fluoride and a protease inhibitor cocktail (pepstatin 1.0 μ g/ml, leupeptin 1.0 μ g/ml, bestatin 1.0 μ g/ml, and aprotinin 1.0 μ g/ml) (Sigma-Aldrich). Cell lysates were centrifuged at 20,000 g for 15 min at 4°C and total protein was estimated using the Bradford assay (BioRad). For Western blot analysis, normalized protein extract was resolved by SDS-PAGE. Proteins were then transferred onto Immobilon P-transfer membranes (Millipore), blocked with TBS-T (25 mM Tris•HCl, pH 7.6, 150 mM NaCl, and 1.0% (v/v) Tween 20) supplemented with 5% (w/v) non-fat dried milk and then incubated with the following antibodies: β -actin (1:20,000; Sigma-Aldrich), CTGF (1:2,000; Santa-Cruz), E-cadherin, (1:1,000; BD Biosciences), N-cadherin, (1:1,000; BD Biosciences) and Jagged1 (1:2,000; Santa-Cruz). Membranes were subsequently incubated with horseradish peroxidase-linked secondary antibodies (New England Biolabs). Blots were developed using enhanced chemiluminescence reagents (SupersignalTM). cDNA was synthesized from 1 μ g of total RNA using the Superscript-II RNase H-Reverse transcriptase cDNA synthesis kit (Invitrogen). Real-Time TaqMan PCR was used to quantify relative gene expression levels of CDH1, CDH2, CTGF, JAG1, SERPINE1, IGFBP3, CGB, SOD2, PLAT and BIRC3 from total RNA extracted from HK-2 cells. The PCR primers and TaqMan probe for all assays were supplied as a pre-optimized single tube primer/probe Gene Expression Assay (Applied Biosystems). 18s rRNA was used as an endogenous control for normalization of the target genes. PCR reactions were set up with Taqman Universal PCR Master Mix in a 10 μ l reaction using 1:20 diluted cDNA. Amplification was performed on the 7900HT Sequence Detection System (Applied Biosystems). Results were analyzed using the Δ Ct method of analysis.

2.3. Library preparation for Illumina sequencing

The Illumina library was prepared according to the manufacturer's instructions. Briefly, we used Dynal oligo (dT) beads (Invitrogen) to isolate poly (A) mRNA from 6 μ g of total RNA. Zinc-mediated fragmentation of mRNA was performed using an RNA fragmentation kit (Ambion), followed by first- and second-strand cDNA synthesis using random hexamer primers. Following cDNA synthesis, an 'end-repair' reaction was performed using a combination of T4 DNA polymerase and E. coli DNA polymerase I Klenow Fragment. A single 'A' base was added to the 3' end of end-repaired fragments by using 3'-to-5' exonuclease, followed by ligation of the Illumina adapter using T4 DNA ligase. Adapter ligated DNA templates at approximately 200 bp were gel purified using a QIAquick Gel Extraction kit (Qiagen) and amplified by PCR. Library quality was assessed on a 2% agarose gel and quantitation was performed using a QubitTM Fluorometer and HS dsDNA kit (Invitrogen).

2.4. Read mapping and statistical analysis

RNA-seq reads were mapped to the human genome (UCSC version hg18) using the Bowtie software [11]. Gene counts were generated from mapped reads using the ERANGE library of python scripts, which incorporates methods for mapping spliced reads and predicting most likely mapping location for reads mapped in multiple locations. Final gene counts were imported into R; differentially expressed genes were identified using the DESeq library of functions. For differential expression analysis, RNA-seq data was filtered to include only genes with at least three reads mapped in at least two lanes, to avoid biologically ambiguous results. P-values were adjusted for multiple testing, using the method proposed by Benjamini and Hochberg [12].

2.5. Promoter analysis

Expression data were first scaled and centered to zero to avoid spurious correlations between strongly differentially expressed genes. Genes were then subdivided into groups using MBC, as implemented using the Mclust library of functions in R. Only genes which were assigned to a single cluster with greater than 90% probability were retained for further analysis. Analysis of promoter sequences in these 584 genes was performed using Genomatix Gene2promoter to extract the 1kb promoters for each group of genes (www.genomatix.de/en/index.html). Each group was tested with Genomatix RegionMiner for overrepresentation of TFBS against a promoter background. Group over- and under-representation of specific TFBS families were hierarchically clustered using Gene Cluster 3.0 (bonsai.hgc.jp/~mdehoon/software/cluster/). Group functional annotations were attributed using the DAVID Bioinformatics Database.

2.6. Statistical analysis of microarray datasets

Detailed methods on the experimental design for the HK-2 microarray are previously published [13]. Published human DN microarray data [10] was downloaded from Nephromine (www.nephromine.org). Array data were first normalized using the *gcrma* library [14] in R (www.cran.org), then filtered to remove genes with mas5 'absent' calls on all arrays, using the *detection.p.val* function in the simpleaffy library [15] in R. The limma library [16] was used to identify differentially expressed genes; p values were adjusted for multiple testing, using the method proposed by Benjamini and Hochberg [12]. Cluster analysis of log transformed normalized expression data was performed using Hierarchical Clustering Explorer (www.cs.umd.edu/hcil/hce/).

2.7. siRNA transfections of renal epithelial cells

siGENOME SMARTpool ARK5/TGFBI siRNA, and siGENOME RISC-Free control siRNA were purchased from Dharmacon. siRNAs were transfected into HK-2 cells at 60%

confluence using Lipofectamine 2000 (Invitrogen) at a final concentration of 20 nM for 24 hr. Cells were then stimulated with TGF- β 1 (5 ng/ml PromoCell) or vehicle for 48 hr.

2.8. qRT-PCR validations in human biopsy samples

Human renal biopsy specimens were procured in an international multicenter study, the European Renal cDNA Bank-Kroener-Fresenius Biopsy bank (ERCB-KFB). Biopsies were obtained from patients after informed consent and with approval of the local ethics committees. Following renal biopsy, the tissue was transferred to RNase inhibitor and microdissected into glomerular and tubular fragments. Total RNA was isolated from microdissected tubulointerstitium as described previously [17]. Reverse transcription and real-time RT-PCR were performed as reported earlier [17]. Pre-developed TaqMan reagents were used for human ANKRD56 (Hs02576315_s1), ENTPD8 (Hs01651150_m1), as well as the reference gene 18S rRNA (Applied Biosystems). The expression of ANKRD56 and ENTPD8 was normalized to 18S rRNA. The mRNA expression was analyzed by standard curve quantification [18]. For the real time RT-PCR data statistical analysis was performed using Kruskal-Wallis and Mann-Whitney U test.

3. RESULTS

3.1. Global expression profile of HK-2 cells treated with TGF- β 1

Immortalized human proximal tubular epithelial (HK-2) cells were used to examine fibrotic responses to TGF- β 1 (5 ng/ml; 48 hr). TGF- β 1 induced a loss of E-cadherin (epithelial marker) and gain of N-cadherin (mesenchymal marker), as observed by Westernblot analysis (Fig. 1A). Increased expression of connective tissue growth factor (CTGF) and the Notch pathway ligand Jagged1 was also observed following TGF- β 1 treatment (Fig. 1A). We have previously shown that Jagged1 expression is elevated in HK-2 cells treated with TGF- β 1, and also in biopsies from DN patients, with co-localization in areas of TIF [13]. Changes in the RNA levels of E-cadherin, N-cadherin, Jagged1 and CTGF reflected altered protein expression in these cells (Fig. 1B).

Having validated TGF- β 1-induced signaling in HK-2 cells, we then examined global gene expression changes using massively parallel sequencing technology (RNA-Seq). 2027 genes were significantly regulated by TGF- β 1 treatment (828 upregulated and 1199 downregulated; Supplementary Table 1 online). Top 20 upregulated and downregulated genes ranked according to p-value are described in Table 1. The top significantly upregulated gene (TGFB1) encodes transforming growth factor, beta-induced protein, an RGD-containing protein that binds to type I, II and IV collagens. Among the top downregulated genes were semaphorin 5B (SEMA5B) and prostaglandin D-synthase (PTGDS). The downregulation of PTGDS is noteworthy given the proposed role of its product prostaglandin D2 in inhibiting TGF- β 1-induced EMT in experimental system [19], and evidence of nephropathy (glomerular hypertrophy, tubular damage and fibrosis) in PTGDS knockout mice [20].

Signaling pathways regulated by TGF- β 1 included: focal adhesion, cell adhesion, ECM-receptor interaction, and complement and coagulation pathways (Table 2). A total of 21 genes within the TGF- β 1 signaling network were differentially expressed, including the induction of TGFB2, TGFB3, TGF- β receptor 1 (TGFB1) and several downstream regulators: Smad ubiquitination regulatory factor 2 (SMURF2), SMADs 4, 7 and 9 (Fig. 2). Enriched gene ontology (GO) cellular compartment categories included genes involved in cytoskeletal reorganization and adhesion, such as basement membrane, extracellular space, integrin complex, focal adhesion and cell-cell adherens junction (Supplementary Table 2 online). Within the RNA-Seq dataset, 134 differentially expressed transcriptional regulators were identified (61 upregulated, 73 downregulated), with several of these upregulated TFs

implicated in fibrotic disorders and EMT: ZEB1, TCF4, SMAD3, SMAD7, ID1, ID3 (Supplementary Table 3 online).

3.2. Promoter analysis – transcriptional network activation

We inferred putative TGF- β 1-driven transcription regulation changes by identifying significantly overrepresented TFBSs among the set of differentially regulated genes using model based clustering (MBC) methods [21]. Using MBC we partitioned differentially expressed genes into 36 independent non-redundant groups and proximal promoter sequence of 1kbp upstream of each gene was retrieved and each group tested for over- and under-representation of TFBSs against a promoter background. Hierarchical clustering based on TFBSs distribution revealed that clustering occurred not according to up or down regulation of expression but more closely by functional annotation (Fig. 3). Clusters split primarily into two segments: the larger cluster (cluster 1) which annotated significantly with pathways relating to focal adhesions, ECM receptor interaction, cell adhesion molecules and other fibrosis related pathways (Supplementary Table 4 online). Within the second, smaller cluster (cluster 2) enriched pathways included MAPK signaling and colorectal cancer, with an absence of functional annotation with fibrosis related pathways observed. Two sets of overrepresented TFBS, containing 8 and 13 TFBS respectively segregated with cluster 1, constituting a cohort of TFs putatively driving the fibrotic signal within our experiment (Fig. 3) (Supplementary Table 5 online). Among these, an enrichment of binding sites for NF κ B, DEAF1 (deformed epidermal autoregulatory factor 1) and E2F7 (E2F transcription factor 7) was evident.

3.3. Comparison of RNA-Seq and microarray data - RNA-Seq identifies a cohort of low abundance transcripts

We previously performed microarray analysis of HK-2 cells treated with TGF- β 1 using an identical design to the present study (GEO: GSE23338). Comparison of mean expression levels of all genes in the RNA-Seq dataset to those determined by Affymetrix microarray (11,033 genes) showed a highly significant signal correlation; Pearson correlation of 0.69, $p < 2.2 \times 10^{-16}$ (Fig. 4A). In a comparison of log fold-change value for each gene, we divided genes into three tertiles of expression level based on the RPKM values (average number of sequence reads per kilo-base pair of coding sequence, per million reads mapped) [22] and found increasing expression correlation for genes with higher expression level (Fig. 4B). In order to determine the extent of additional information available from RNA-Seq data, we compared the differentially expressed genes in both datasets (microarray – 199; RNA-Seq – 2027) (Fig. 4C). Among the 352 low expression genes detected only by RNA-Seq are 29 involved in cell adhesion – including claudin 9, collagen type-alpha 1, integrin alpha M and pleckstrin (Supplementary Table 6 online). Similarly, in comparing RNA-Seq and microarray on the basis of the genes present in the array data set (8632 genes) we observed greater overlap at higher expression level – and consistently higher sensitivity in detection of differentially expressed genes on the RNA-Seq platform (1517 genes with RNA-Seq; 199 with microarray). We then compared significantly enriched pathways ($p < 0.05$) that were detected on each platform (Fig. 4D). RNA-Seq detected four times more significantly regulated pathways, and among the pathways detected by RNA-Seq only were those known to play an important role in renal fibrosis, such as focal adhesion, cell adhesion molecules and ECM receptor interaction.

In order to validate findings from the sequence data, six genes were selected for qRT-PCR analysis based on their respective RNA-Seq RPKM expression values (high, medium, low expressors). Three up-regulated genes – SERPINE1 (high), IGFBP3 (medium), CGB5 (low); and three down-regulated genes – SOD2 (high), PLAT (medium), and BIRC3 (low) were assessed. All genes were successfully validated, with TGF- β 1 treatment inducing

SERPINE1, IGFBP3 and CGB, and repressing SOD2, PLAT and BIRC3 expression (Fig. 4E). Comparison of qRT-PCR fold-changes with RNA-Seq and microarray data indicated that for the six selected targets, RNA-Seq data best correlated with real-time validations.

3.4. Intersection of in vitro data with human DN expression data

In order to extrapolate sequence data from the HK-2 cell model to human DN, we compared our *in vitro* sequence data to available human tubulo-interstitial array data of DN patients (n=11) compared to minimal change disease patients (MCD; n=4) and controls (living donors; n=3) [10]. Of the 2027 genes identified by RNA-Seq as significantly regulated by TGF- β 1, corresponding microarray expression data was available for 1356 genes (Fig. 5A). A total of 179 genes were significantly differentially expressed in both the RNA-Seq and human DN datasets (Fig. 5B) (Supplementary Table 7 online). Gene ontology analysis of the 179 genes indicated an enrichment of ECM and cytoskeleton-associated gene products (Supplementary Table 8 online). Upregulated genes included: collagen family members (COL4A1, COL4A2, COL16A1); key markers of EMT and TIF (VIM, THBS1, ITGAV); and the Notch pathway ligand JAGGED1.

Closer examination of known early phase EMT TFs and adhesion molecules revealed expression variation in SNAI2, CLDN4 and CLDN3 that was only evident in mild DN relative to healthy donor (Supplementary Figure 1 online). The SNAIL family of TFs are induced in response to TGF- β 1, in a SMAD-dependent manner, and act to repress epithelial marker gene expression (including claudins), and activate mesenchymal markers [23]. Expression of SNAI2 appeared to be induced in mild DN, and restored to the level of healthy donors in moderate and severe DN; the opposite pattern was noted in CLDN3 and CLDN4. These changes were moderately significant at nominal $p=0.037$, 0.016 and 0.027 for SNAI2, CLDN3 and CLDN4, respectively. Similar patterns of change were observed in sequence data in response to TGF- β 1 for SNAI2 and CLDN3 (nominal $p=2.83 \times 10^{-5}$ and 2.45×10^{-5}). Additional cell adhesion-related genes showing significant transient downregulation in mild DN included LAMC2, RHOB, CDH19 and SELE (data not shown).

3.5. TGFBI and ARK5: functional role in TGF- β 1-driven fibrotic signaling in renal epithelial cells

Among the genes upregulated by TGF- β 1 and also elevated in human DN was ARK5, encoding a serine/threonine AMPK-related kinase previously reported to be a key player in tumor invasion and directly activated by Akt [24]. To demonstrate the biological significance of the RNA-Seq data, the effect of siRNA-mediated inhibition of these genes on EMT and TGF- β 1 signaling was assessed. Inhibition of ARK5 significantly attenuated TGF- β 1-induced Jagged1 and N-cadherin RNA expression towards basal expression levels, with no effect seen on E-cadherin or CTGF expression. These data were subsequently validated at the protein level (Fig. 6, A and B). TGFBI was the most significantly upregulated gene in response to TGF- β 1 and is also elevated in human DN. Silencing of TGFBI resulted in an induction of basal E-cadherin expression, evident at both the RNA and protein level (Fig. 7, A and B), suggesting that TGFBI negatively regulates E-cadherin expression in renal epithelial cells. The observed induction of E-cadherin expression was reversed in the presence of TGF- β 1, suggesting that while TGFBI may negatively regulate E-cadherin expression, it is not essential for TGF- β 1-induced loss of E-cadherin.

3.6. ANKRD56 and ENTPD8 are low abundance transcripts downregulated in renal epithelial cells by TGF- β 1 and also in human DN renal biopsies

In order to determine whether RNA-Seq data from our cell model could inform the selection of novel low expression genes in human DN, previously undetectable by microarray, we selected a cohort of low abundance transcripts from the RNA-Seq dataset and validated

these by qRT-PCR in human DN samples (n=14) versus living donor controls (n=7). Expression changes in five genes were measured - ANXA8, ANKRD56, ENTPD8, TSHZ3 and NTF6A. For three genes (ANXA8, TSHZ3, NTF6A), the low overall expression levels were below the threshold of detection by qRT-PCR. ANKRD56 and ENTPD8 expression levels were significantly lower in DN biopsies, mimicking the expression response observed in TGF- β 1 stimulated HK-2 cells (Fig. 8).

4. Discussion

TGF- β 1 is implicated in multiple fibrotic disorders, and is recognized as a major driver of renal fibrosis including renal tubule injury. TGF- β 1 stimulation of cultured renal epithelia is a conventional model of DN associated TIF [25–28]. However the full extent of transcriptomic response to TGF- β 1 treatment has not been clarified. Using RNA-Seq we have performed detailed analyses of transcriptional changes induced by pro-fibrotic TGF- β 1 in renal tubule epithelial cells. By comparing RNA-Seq and microarray data using identical experimental conditions, we demonstrate that while there is a high degree of correlation between both platforms, RNA-Seq uniquely identifies a large cohort of differentially expressed low abundance transcripts. By comparing our *in vitro* model RNA-Seq data with array data from human DN renal biopsies we have: defined a subset of 179 co-regulated genes, which may represent the TGF- β 1-driven signature in renal fibrosis; identified ARK5 and TGFBI as putative mediators of TGF- β 1 signaling in renal epithelial cells; and identified the low abundance transcripts ENTPD8 and ANKRD56 as differentially expressed in human DN renal biopsies. These data provide a comprehensive overview of transcriptomic responses of renal epithelia exposed to TGF- β 1 and define those genes most relevant in the context of DN.

By examining the promoter regions of TGF- β 1 responsive genes we have defined a series of potential transcriptional regulators. Among these, an enrichment of binding sites for NF κ B was evident. Several studies have previously reported the activation of the NF κ B signaling pathway in progressive DN [10, 29]. This would suggest that the global transcriptional response in the *in vitro* dataset mimics the response seen in human DN. Though increased transcription factor expression is not a functional necessity for increased transcription factor activity (i.e. increased target expression), a number of these TFs such as KLFs, GLI1, EGR1 and NF κ B2 were also found themselves to be differentially regulated in the RNA-Seq dataset.

We have shown that, as expected, sequencing technology offers substantially enhanced sensitivity in detecting differentially expressed genes when compared with microarrays. Several studies have reported the advantages of sequencing technology versus array data [30, 31]. In our model system, RNA-Seq detected approximately 10 times more differentially expressed genes than array technology, and as a consequence three times more differentially regulated pathways. The lesser degree of correlation in low expression genes is likely explained by non-specific background hybridization on the array [32].

To investigate the relevance of this transcriptomic signature in the context of human DN, we compared sequence data to expression data from human kidney biopsies enriched for tubulointerstitium from patients at varying stages of DN. In this analysis we found a signature of 179 genes commonly regulated in the human and *in vitro* models. We investigated the functional significance of two of these commonly upregulated genes: ARK5 and TGFBI. Suppression of ARK5 expression in tubular epithelial cells attenuated TGF- β 1 signaling. ARK5 encodes an AMPK-related serine/threonine kinase implicated in metastasis and acts downstream of Akt [24, 33]. We have previously detected increased levels of TGF- β 1 signaling and PKB/Akt activity in the kidney tubules of a Goto-Kakizaki rat model of

type 2 diabetes, suggesting that this signaling cascade may be involved in DN-associated EMT *in vivo* [34]. By silencing ARK5 expression, TGF- β 1-driven gain of N-cadherin was impeded. A putative role for ARK5 in mediating TGF- β 1-driven Notch pathway expression is also evident, with ARK5 inhibition preventing the induction of Notch ligand - Jagged1. Expression of the Notch pathway ligand Jagged1 was elevated in human DN, and we have recently shown that several genes involved in the Notch pathway are differentially expressed in renal biopsies from individuals with DN [13]. TGFBI encodes an ECM protein induced by TGF- β 1 that is associated with other ECM proteins and functions as a ligand for integrins [35]. Inhibition of TGFBI has been shown to significantly reduce metastasis in colon cancer [36]. Here we demonstrate that basal expression of E-cadherin was significantly upregulated upon TGFBI silencing, suggesting this protein may act as a negative regulator of E-cadherin expression. The fact that TGF- β 1 still suppressed E-cadherin levels in TGFBI silenced cells would indicate that TGFBI is not an absolute requirement.

Examination of expression changes of an early phase regulator of EMT, the transcriptional repressor SNAI2, and SNAI2 targets (CLDN3, CLDN4) [23, 37], revealed that SNAI2 was upregulated in mild DN patients versus healthy donor, and subsequently restored to basal levels in moderate and severe DN patient. A corresponding repression in CLDN3 and CLDN4 expression was observed in mild DN patients, which was subsequently restored in moderate and severe DN patients. Comparison of these data with our *in vitro* model data revealed that the expression changes observed in mild DN best matched the observed response to TGF- β 1 in renal epithelial cells.

The higher degree of sensitivity offered by RNA-Seq allows for the identification of low abundance transcripts that are undetectable by microarray [38, 39]. Using the additional information we acquired from sequence data on low abundance transcripts regulated by TGF- β 1, which were previously not available for measure on our microarray panels, we selected several of these and measured mRNA levels in renal biopsy material from DN patients versus living donor controls. Of the five transcripts measured, two of these were detected (ENTPD8 and ANKRD56) and validated the TGF- β 1 response in renal epithelial cell sequence data. In human DN we observed a downregulation of ENTPD8 expression, which encodes an intracellular enzyme that regulates the availability of extracellular nucleotide levels (ATP, ADP) to receptors [40]. Aberrant nucleotide signaling has previously been shown to promote injury in DN, with another family member, ENTPD1, shown to be a protective factor in DN [41]. ANKRD56 was also downregulated in human DN in the present study, and EST database analysis of both human and mouse genomes shows that ANKRD56 RNA is present in kidney, intestine, mammary gland, muscle and uterus <http://www.ncbi.nlm.nih.gov/UniGene/clust.cgi?ORG=Hs&CID=257292>. ANKRD56 contains multiple ankyrin repeats consisting of 30–34 amino acid residues, which mediate protein-protein interactions[42]. Recently, inactivation of ANKRD26 was shown to induce obesity and diabetes in mice, hinting at a putative role for these proteins in disease [43, 44]. Currently the precise roles of ENTPD8 and ANKRD56 remain undefined and the function of these genes in the context of DN needs to be addressed.

We recognize several limitations to the present study which need to be considered when interpreting these results. While we have compared our *in vitro* model RNA-seq dataset with a human disease microarray dataset, a comparison on identical platforms (i.e. RNA-seq on human DN biopsies) would be a more powerful approach to identify low abundance transcripts. However, presently no such suitable dataset is available. Furthermore, the human disease dataset employed in this study is limited with respect to sample size and clinical parameter information necessary to assess a correlation between differentially expressed transcripts and markers of disease progression. Finally, comparison of the cell model dataset with *in vivo* expression datasets is an attractive approach which we intend to develop in

future studies. It is noteworthy that as yet no RNA-seq datasets are available from *in vivo* models of DN, notwithstanding the controversies about such models [45].

In summary, using tubular epithelial cells stimulated with the pro-fibrotic cytokine TGF- β 1 as a model of renal fibrosis, we have performed a comprehensive analysis of the transcriptome response using both sequence and array platforms. In this study we: (1) identified a cluster of fibrosis-related genes predicted to be regulated by a relatively small group of TFs, including both established and novel regulators of EMT. These TFs may act as master regulators of fibrosis in this cell model and are worthy of follow up functional studies; (2) we have jointly analyzed our *in vitro* data with human DN biopsy data in order to define a co-regulated subset of genes. These genes define the responsive subset in human DN to TGF- β 1, a key driver of renal fibrosis; and finally (3) using RNA-Seq data from our renal epithelial cell model system, we have identified novel low abundance transcripts downregulated in human DN, providing evidence that not all DN-associated transcripts have been identified. Targeted intervention of these genes may offer novel therapeutic approaches in an effort to attenuate the initiation and progression of renal fibrosis.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Dr. Amanda Lohan for technical advice on library preparation. EPB is supported by a Science Foundation Ireland US-Ireland R&D Partnership and a Science Foundation Ireland North South Research Partnership. SAR was supported by an Irish Health Research Board Clinical Fellowship. DPB is supported by DEL Northern Ireland and Action Medical Research UK. CG is a Science Foundation Ireland Principal Investigator.

References

- Collins AJ, Foley RN, Herzog C, Chavers B, Gilbertson D, Ishani A, Kasiske B, Liu J, Mau LW, McBean M, Murray A, St Peter W, Guo H, Gustafson S, Li Q, Li S, Peng Y, Qiu Y, Roberts T, Skeans M, Snyder J, Solid C, Wang C, Weinhandl E, Zaun D, Arko C, Chen SC, Dalleska F, Daniels F, Dunning S, Ebben J, Frazier E, Hanzlik C, Johnson R, Sheets D, Wang X, Forrest B, Constantini E, Everson S, Eggers P, Agodoa L. US Renal Data System 2010 Annual Data Report. *Am J Kidney Dis.* 2011; 57:A8, e1–526. [PubMed: 21184928]
- Reeves WB, Andreoli TE. Transforming growth factor beta contributes to progressive diabetic nephropathy. *Proc Natl Acad Sci U S A.* 2000; 97:7667–7669. [PubMed: 10884396]
- Kalluri R, Neilson EG. Epithelial-mesenchymal transition and its implications for fibrosis. *J Clin Invest.* 2003; 112:1776–1784. [PubMed: 14679171]
- Russo LM, del Re E, Brown D, Lin HY. Evidence for a role of transforming growth factor (TGF)-beta1 in the induction of postglomerular albuminuria in diabetic nephropathy: amelioration by soluble TGF-beta type II receptor. *Diabetes.* 2007; 56:380–388. [PubMed: 17259382]
- Shi Y, Massague J. Mechanisms of TGF-beta signaling from cell membrane to the nucleus. *Cell.* 2003; 113:685–700. [PubMed: 12809600]
- Schmierer B, Hill CS. TGFbeta-SMAD signal transduction: molecular specificity and functional flexibility. *Nat Rev Mol Cell Biol.* 2007; 8:970–982. [PubMed: 18000526]
- Chaudhury A, Howe PH. The tale of transforming growth factor-beta (TGFbeta) signaling: a soigne enigma. *IUBMB Life.* 2009; 61:929–939. [PubMed: 19787707]
- Kalluri R, Weinberg RA. The basics of epithelial-mesenchymal transition. *J Clin Invest.* 2009; 119:1420–1428. [PubMed: 19487818]
- Canales RD, Luo Y, Willey JC, Austermilller B, Barbacioru CC, Boysen C, Hunkapiller K, Jensen RV, Knight CR, Lee KY, Ma Y, Maqsoodi B, Papallo A, Peters EH, Poulter K, Ruppel PL, Samaha RR, Shi L, Yang W, Zhang L, Goodsaid FM. Evaluation of DNA microarray results with quantitative gene expression platforms. *Nat Biotechnol.* 2006; 24:1115–1122. [PubMed: 16964225]

10. Schmid H, Boucherot A, Yasuda Y, Henger A, Brunner B, Eichinger F, Nitsche A, Kiss E, Bleich M, Grone HJ, Nelson PJ, Schlondorff D, Cohen CD, Kretzler M. Modular activation of nuclear factor-kappaB transcriptional programs in human diabetic nephropathy. *Diabetes*. 2006; 55:2993–3003. [PubMed: 17065335]
11. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol*. 2009; 10:R25. [PubMed: 19261174]
12. Benjamini Y, Drai D, Elmer G, Kafkafi N, Golani I. Controlling the false discovery rate in behavior genetics research. *Behav Brain Res*. 2001; 125:279–284. [PubMed: 11682119]
13. Walsh DW, Roxburgh SA, McGettigan P, Berthier CC, Higgins DG, Kretzler M, Cohen CD, Mezzano S, Brazil DP, Martin F. Co-regulation of Gremlin and Notch signalling in diabetic nephropathy. *Biochim Biophys Acta*. 2008; 1782:10–21. [PubMed: 17980714]
14. Wu, RAI Zhijin; Gentleman, Robert; Martinez-Murillo, Francisco; Spencer, Forrest. A Model-Based Background Adjustment for Oligonucleotide Expression Arrays. *Journal of the American Statistical Association*. 2004; 99:9.
15. Wilson CL, Miller CJ. Simpleaffy: a BioConductor package for Affymetrix Quality Control and data analysis. *Bioinformatics*. 2005; 21:3683–3685. [PubMed: 16076888]
16. Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol*. 2004; 3:Article3. [PubMed: 16646809]
17. Cohen CD, Frach K, Schlondorff D, Kretzler M. Quantitative gene expression analysis in renal biopsies: a novel protocol for a high-throughput multicenter application. *Kidney Int*. 2002; 61:133–140. [PubMed: 11786093]
18. Schmid H, Henger A, Cohen CD, Frach K, Grone HJ, Schlondorff D, Kretzler M. Gene expression profiles of podocyte-associated molecules as diagnostic markers in acquired proteinuric diseases. *J Am Soc Nephrol*. 2003; 14:2958–2966. [PubMed: 14569107]
19. Zhang A, Dong Z, Yang T. Prostaglandin D2 inhibits TGF-beta1-induced epithelial-to-mesenchymal transition in MDCK cells. *Am J Physiol Renal Physiol*. 2006; 291:F1332–1342. [PubMed: 16896186]
20. Ragolia L, Palaia T, Hall CE, Maesaka JK, Eguchi N, Urade Y. Accelerated glucose intolerance, nephropathy, and atherosclerosis in prostaglandin D2 synthase knock-out mice. *J Biol Chem*. 2005; 280:29946–29955. [PubMed: 15970590]
21. Fraley C, Raftery A. Model-based clustering, Discriminant analysis, and Density Estimation. *Journal of the American Statistical Association*. 2002; 97:611–631.
22. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods*. 2008; 5:621–628. [PubMed: 18516045]
23. Martinez-Estrada OM, Culleres A, Soriano FX, Peinado H, Bolos V, Martinez FO, Reina M, Cano A, Fabre M, Vilaro S. The transcription factors Slug and Snail act as repressors of Claudin-1 expression in epithelial cells. *Biochem J*. 2006; 394:449–457. [PubMed: 16232121]
24. Suzuki A, Lu J, Kusakai G, Kishimoto A, Ogura T, Esumi H. ARK5 is a tumor invasion-associated factor downstream of Akt signaling. *Mol Cell Biol*. 2004; 24:3526–3535. [PubMed: 15060171]
25. Nyhan KC, Faherty N, Murray G, Cooley LB, Godson C, Crean JK, Brazil DP. Jagged/Notch signalling is required for a subset of TGFbeta1 responses in human kidney epithelial cells. *Biochim Biophys Acta*. 2010; 1803:1386–1395. [PubMed: 20833210]
26. Lin H, Wang D, Wu T, Dong C, Shen N, Sun Y, Xie H, Wang N, Shan L. Blocking core fucosylation of TGF-beta1 receptors downregulates their functions and attenuates the epithelial-mesenchymal transition of renal tubular cells. *Am J Physiol Renal Physiol*. 2011; 300:F1017–1025. [PubMed: 21228108]
27. Xu Y, Wan J, Jiang D, Wu X. BMP-7 counteracts TGF-beta1-induced epithelial-to-mesenchymal transition in human renal proximal tubular epithelial cells. *J Nephrol*. 2009; 22:403–410. [PubMed: 19557718]
28. Zeisberg M, Hanai J, Sugimoto H, Mammoto T, Charytan D, Strutz F, Kalluri R. BMP-7 counteracts TGF-beta1-induced epithelial-to-mesenchymal transition and reverses chronic renal injury. *Nat Med*. 2003; 9:964–968. [PubMed: 12808448]

29. Mezzano S, Aros C, Droguett A, Burgos ME, Ardiles L, Flores C, Schneider H, Ruiz-Ortega M, Egido J. NF-kappaB activation and overexpression of regulated genes in human diabetic nephropathy. *Nephrol Dial Transplant*. 2004; 19:2505–2512. [PubMed: 15280531]
30. Marioni JC, Mason CE, Mane SM, Stephens M, Gilad Y. RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. *Genome Res*. 2008; 18:1509–1517. [PubMed: 18550803]
31. Bottomly D, Walter NA, Hunter JE, Darakjian P, Kawane S, Buck KJ, Searles RP, Mooney M, McWeeney SK, Hitzemann R. Evaluating gene expression in C57BL/6J and DBA/2J mouse striatum using RNA-Seq and microarrays. *PLoS One*. 2011; 6:e17820. [PubMed: 21455293]
32. Ramdas L, Cogdell DE, Jia JY, Taylor EE, Dunmire VR, Hu L, Hamilton SR, Zhang W. Improving signal intensities for genes with low-expression on oligonucleotide microarrays. *BMC Genomics*. 2004; 5:35. [PubMed: 15196312]
33. Suzuki A, Iida S, Kato-Uranishi M, Tajima E, Zhan F, Hanamura I, Huang Y, Ogura T, Takahashi S, Ueda R, Barlogie B, Shaughnessy J Jr, Esumi H. ARK5 is transcriptionally regulated by the Large-MAF family and mediates IGF-1-induced cell invasion in multiple myeloma: ARK5 as a new molecular determinant of malignant multiple myeloma. *Oncogene*. 2005; 24:6936–6944. [PubMed: 16044163]
34. Kattla JJ, Carew RM, Heljic M, Godson C, Brazil DP. Protein kinase B/Akt activity is involved in renal TGF-beta1-driven epithelial-mesenchymal transition in vitro and in vivo. *Am J Physiol Renal Physiol*. 2008; 295:F215–225. [PubMed: 18495798]
35. Kim JE, Kim SJ, Lee BH, Park RW, Kim KS, Kim IS. Identification of motifs for cell adhesion within the repeated domains of transforming growth factor-beta-induced gene, betaig-h3. *J Biol Chem*. 2000; 275:30907–30915. [PubMed: 10906123]
36. Ma C, Rong Y, Radloff DR, Datto MB, Centeno B, Bao S, Cheng AW, Lin F, Jiang S, Yeatman TJ, Wang XF. Extracellular matrix protein betaig-h3/TGFBI promotes metastasis of colon cancer by enhancing cell extravasation. *Genes Dev*. 2008; 22:308–321. [PubMed: 18245446]
37. Kajita M, McClintock KN, Wade PA. Aberrant expression of the transcription factors snail and slug alters the response to genotoxic stress. *Mol Cell Biol*. 2004; 24:7559–7566. [PubMed: 15314165]
38. Nagalakshmi U, Waern K, Snyder M. RNA-Seq: a method for comprehensive transcriptome analysis. *Curr Protoc Mol Biol*. 2010; Chapter 4(Unit 4 11):11–13. [PubMed: 20069539]
39. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet*. 2009; 10:57–63. [PubMed: 19015660]
40. Shirley DG, Vekaria RM, Sevigny J. Ectonucleotidases in the kidney. *Purinergic Signal*. 2009; 5:501–511. [PubMed: 19333785]
41. Friedman DJ, Rennke HG, Csizmadia E, Enjyoji K, Robson SC. The vascular ectonucleotidase ENTPD1 is a novel renoprotective factor in diabetic nephropathy. *Diabetes*. 2007; 56:2371–2379. [PubMed: 17473221]
42. Li J, Mahajan A, Tsai MD. Ankyrin repeat: a unique motif mediating protein-protein interactions. *Biochemistry*. 2006; 45:15168–15178. [PubMed: 17176038]
43. Bera TK, Liu XF, Yamada M, Gavrilova O, Mezey E, Tessarollo L, Anver M, Hahn Y, Lee B, Pastan I. A model for obesity and gigantism due to disruption of the Ankrd26 gene. *Proc Natl Acad Sci U S A*. 2008; 105:270–275. [PubMed: 18162531]
44. Raciti GA, Bera TK, Gavrilova O, Pastan I. Partial inactivation of Ankrd26 causes diabetes with enhanced insulin responsiveness of adipose tissue in mice. *Diabetologia*. 2011
45. Brosius FC 3rd, Alpers CE, Bottinger EP, Breyer MD, Coffman TM, Gurley SB, Harris RC, Kakoki M, Kretzler M, Leiter EH, Levi M, McIndoe RA, Sharma K, Smithies O, Susztak K, Takahashi N, Takahashi T. Mouse models of diabetic nephropathy. *J Am Soc Nephrol*. 2009; 20:2503–2512. [PubMed: 19729434]

Highlights

- TGF- β 1 is implicated in renal tubule injury in diabetic nephropathy (DN).
- RNA-seq identifies a TGF- β 1-driven signal in renal tubule epithelial cells.
- ARK5 is upregulated by TGF- β 1 and necessary for complete TGF- β 1 signaling.
- Defining TGF- β 1 signaling in renal epithelials will aid our understanding of DN.

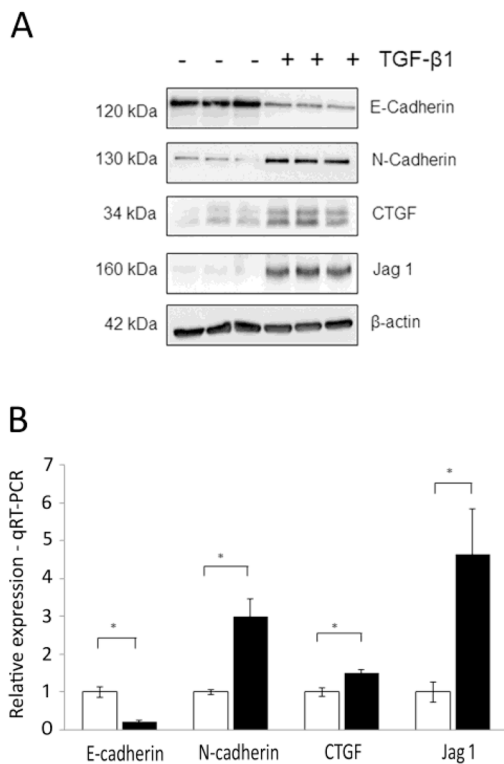


Fig. 1. Establishment of fibrotic phenotype in human proximal tubular epithelial cells. **(A)** Western blot of markers of fibrosis (E-cadherin, N-cadherin) and TGF- β 1 signaling (CTGF, Jagged1) in HK-2 cells stimulated with TGF- β 1 (5ng/ml) for 48hr versus unstimulated controls (n = 3). **(B)** Real-time quantitative PCR validation of differentially expressed transcripts from RNA-Seq experiment. Quantitative TaqMan PCR was performed in unstimulated HK-2 cells (grey bar) and HK-2s stimulated with TGF- β 1 (black bars). Expression was normalized to 18S. Data are plotted as mean \pm SE. * P < 0.05, Student's unpaired t test, n = 3 per group.

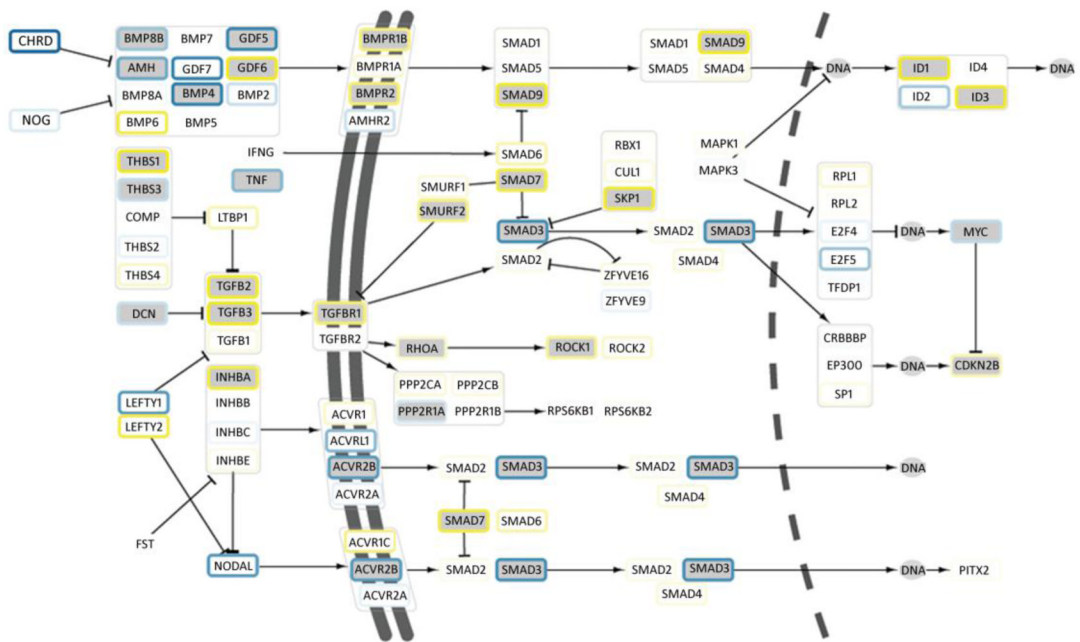


Fig. 2. Transcriptional changes in TGF- β 1 signaling pathway. RNA-Seq data are mapped to the pathway; green nodes show differentially expressed genes. Log-fold change in response to TGF- β 1 treatment is illustrated in node border colours ranging from blue (down-regulated) to yellow (up-regulated). Double and dashed lines represent cell and nuclear membrane. Nodes marked 'DNA' represent transcriptional regulatory interactions.

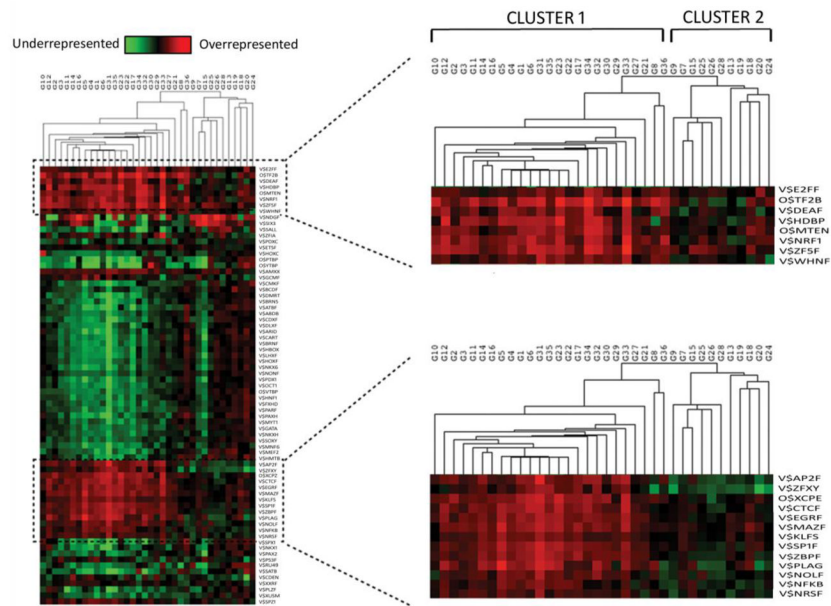
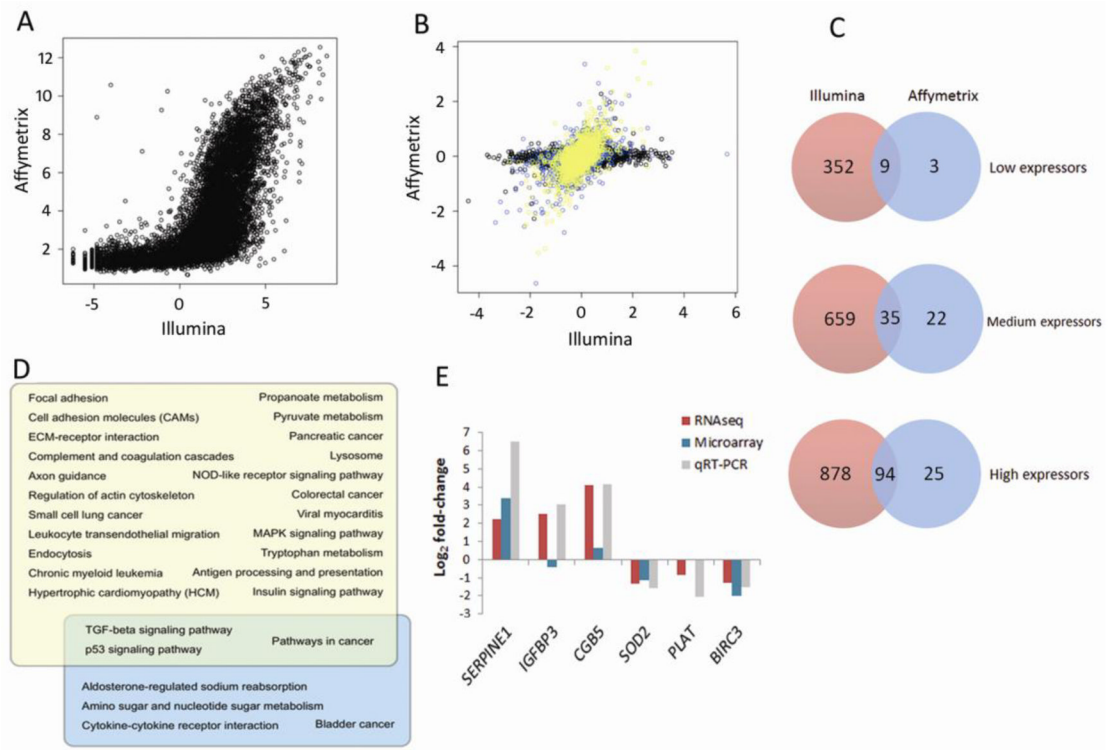


Fig. 3. Analysis of over-represented TFBS among clusters of significantly regulated genes. Differentially expressed genes in RNA-Seq were separated into 36 clusters (G1–36). 1Kbp promoter regions were assessed for overrepresented (red) and underrepresented (green) TFBS in each group.

**Fig. 4.**

Comparison of RNA-Seq to microarray of TGF- β 1 effects on gene expression in HK-2 cell line. **(A)** Raw signal comparison. In left panel, mean RNA-Seq RPKM of each gene is compared to expression level on the microarray (calculated from hybridization intensity). **(B)** Comparison of log FC of each gene from RNA-Seq to microarray. Genes were classified as high (yellow), medium (blue), or low (black) expressors, based on RNA-Seq RPKM for each gene. **(C)** Degree of overlap of differentially expressed genes from RNA-Seq and microarray, subdivided by expression level (RNA-Seq RPKM value; microarray hybridization intensity). **(D)** Overlap of significantly regulated KEGG pathways from RNA-Seq (yellow) and microarray (blue). **(E)** Log-fold changes of six selected genes in RNA-Seq (purple bar), microarray (blue bar) and corresponding quantitative TaqMan PCR validation (grey) in HK-2s stimulated with TGF- β 1 (5ng/ml; 48hr). Expression was normalized to 18S. Data are plotted as mean \pm SE. * $P < 0.05$, Student's unpaired t test, $n = 3$ per group.

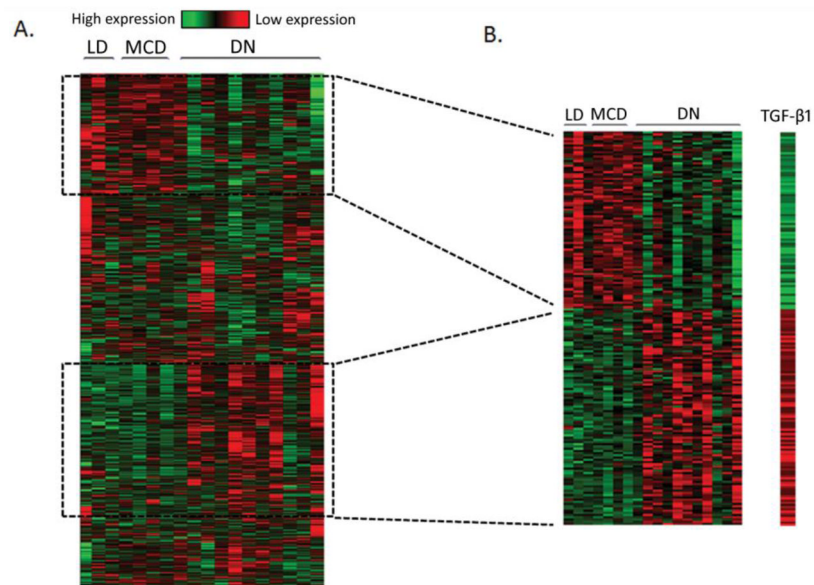


Fig. 5. RNA-Seq identifies a TGF- β 1 regulated gene set with overlapping expression profile in human DN. **(A)** 2027 genes were identified as significantly differentially expressed in renal epithelial cells in response to TGF- β 1 treatment. Heatmap of gene expression patterns of 1356/2027 genes available from microarray data of renal biopsies from human diabetic nephropathy (DN; n=11) versus living donor (LD; n=3) and minimal change disease (MCD; n=4) controls. **(B)** A subset of 179 genes were significantly differentially expressed in both RNA-Seq and human DN microarray datasets. Heatmap of log-fold changes of 179 genes in HK-2 cells stimulated with TGF- β 1 (5ng/ml; 48hr), and gene expression patterns of corresponding 179 genes from microarray data of renal biopsies from human diabetic nephropathy (DN; n=11) versus living donor (LD; n=3) and minimal change disease (MCD; n=4) controls. Microarray data are \log_2 median-centred intensities. Green – downregulated gene expression; Red – upregulated gene expression.

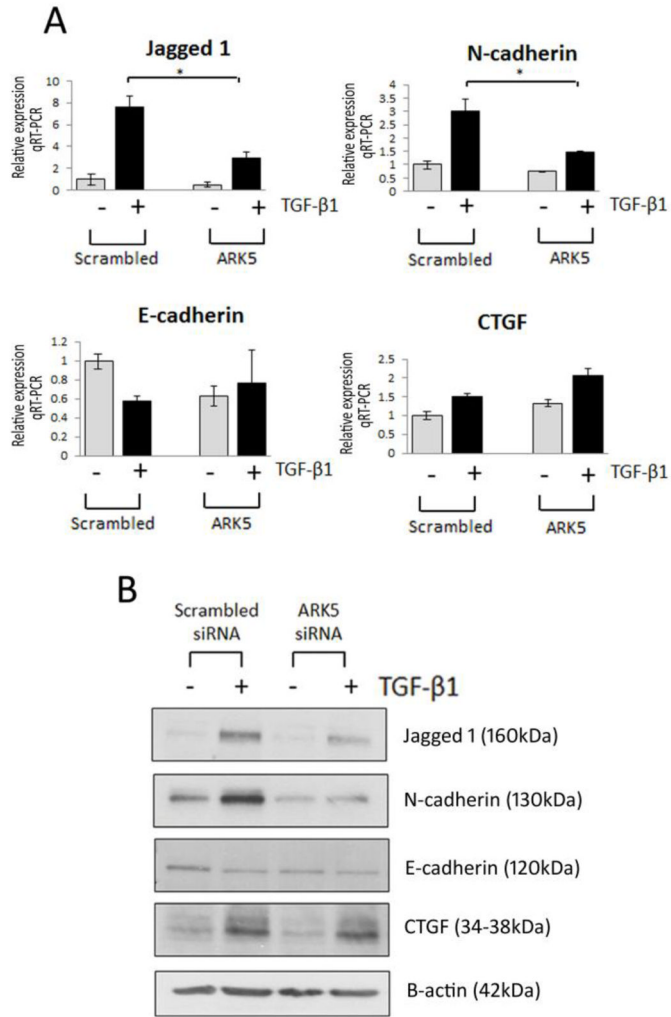


Fig. 6. Role for ARK5 in TGF-β1-induced fibrosis in renal epithelial cells. **(A)** TaqMan quantitative PCR analysis of Jagged1, N-cadherin (CDH2), E-cadherin (CDH1) and CTGF in HK-2 cells transfected with ARK5 siRNA in the absence (grey bar)/presence (black bar) of TGF-β1 (5ng/ml; 48h). **(B)** Representative Western blot of Jagged1, N-cadherin (CDH2), E-cadherin (CDH1) and CTGF HK-2 cells transfected with ARK5 siRNA, respectively. HK-2 cells transfected with scrambled siRNA were selected as a control. For TaqMan PCR, expression was normalized to GAPDH. Data are plotted as mean ± SE. **P* < 0.05, Student's unpaired *t* test, n = 3 per group.

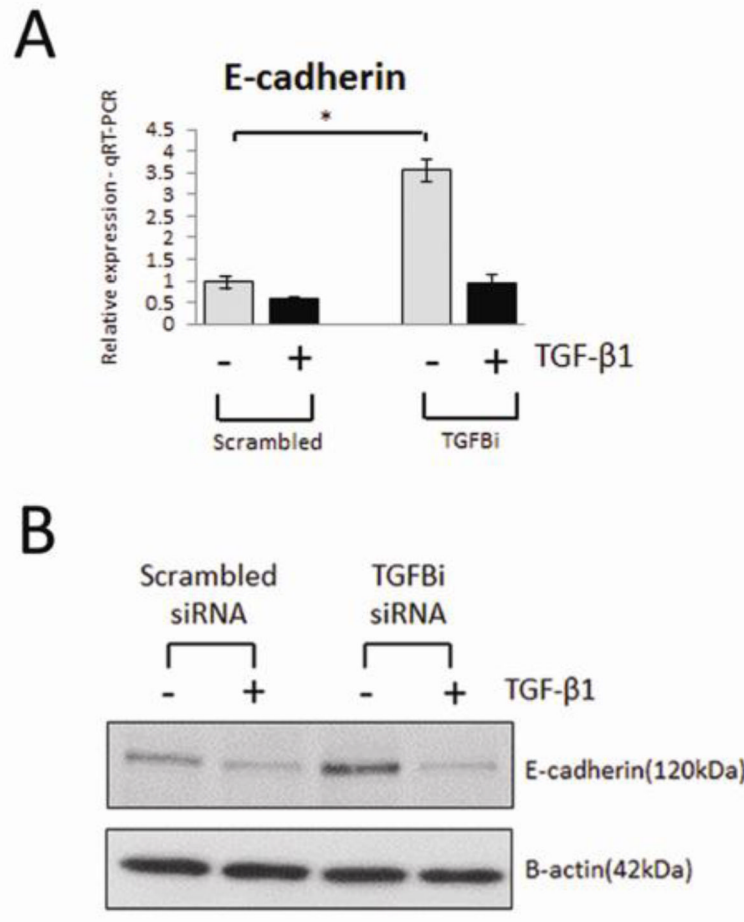


Fig. 7. Silencing of TGFBI induces E-cadherin expression in renal epithelial cells. **(A)** TaqMan quantitative PCR analysis of E-cadherin (CDH1) expression in HK-2 cells transfected with TGFBI siRNA in the absence (grey bar)/presence (black bar) of TGF-β1 (5ng/ml; 48h). **(B)** Representative Western blot of E-cadherin (CDH1) expression in HK-2 cells transfected with TGFBI siRNA. HK-2 cells transfected with scrambled siRNA were selected as a control. For TaqMan PCR, expression was normalized to GAPDH. Data are plotted as mean ± SE. **P* < 0.05, Student's unpaired *t* test, n = 3 per group.

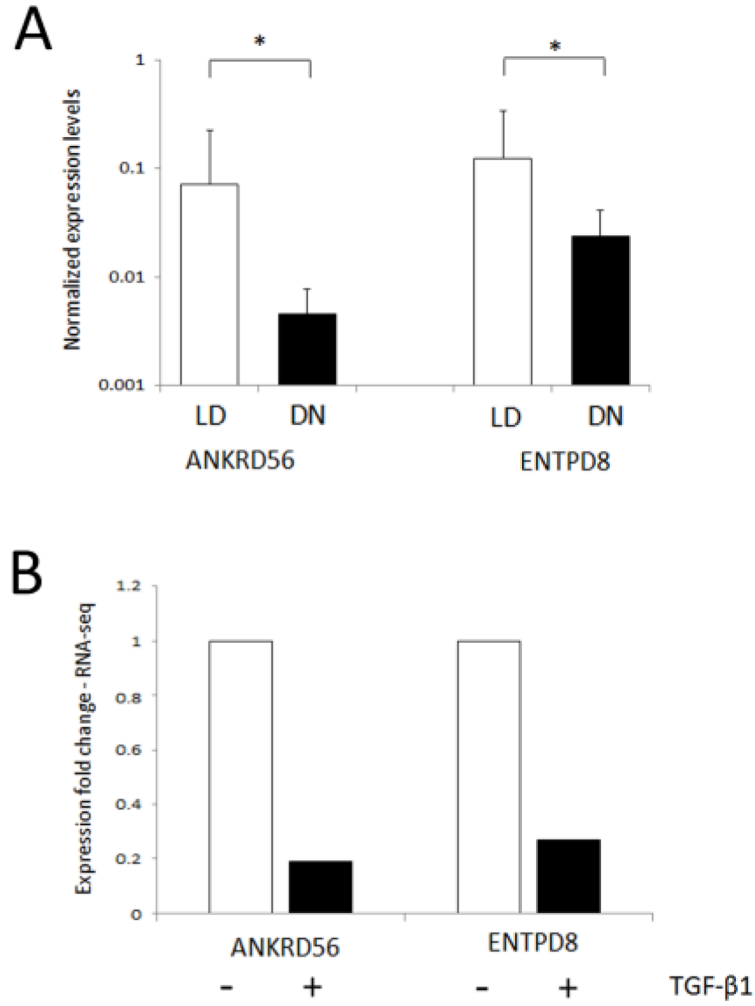


Fig. 8. ANKRD56 and ENTPD8 are differentially expressed in human DN. **(A)** TaqMan quantitative PCR analysis of ANKRD56 and ENTPD8 in human DN renal biopsy samples (DN n=13; black) versus living donor controls (LD: n=7; grey). Data are plotted as mean \pm SE. The graphs for ANKRD56 and ENTPD8 show expression ratios of each gene normalized to 18S rRNA. Statistical analysis was performed using Kruskal-Wallis and Mann-Whitney U test (mean \pm SE. * $P < 0.05$) **(B)** RNA-Seq data fold-change for ANKRD56 and ENTPD8 in HK-2 cells stimulated with (black)/without (grey) TGF- β 1.

Table 1Top 20 significantly upregulated/downregulated genes in HK-2 cells stimulated with TGF- β 1

Gene	Reads *	Log2 Fold Change	Adjusted P-value
TGFB1	10928.30	2.67	3.40E-199
IGFBP3	1290.22	2.59	3.23E-198
COL1A1	1344.63	2.42	3.45E-174
SLCO2A1	752.95	2.21	2.97E-107
SERPINE1	13607.37	2.44	3.14E-101
IL11	538.53	2.32	9.30E-96
CGB8	71.08	4.98	1.72E-76
C20orf58	333.50	1.80	1.37E-57
TMEPAI	2268.56	1.95	2.61E-54
C4orf26	216.30	2.27	3.81E-54
JAG1	784.27	2.12	3.22E-50
ADAM12	534.83	1.71	5.37E-47
TAGLN	3525.86	1.75	1.35E-46
GDF6	100.20	2.92	1.97E-46
CDH2	1812.90	1.54	1.27E-44
CGB5	31.53	4.66	5.27E-40
KRT7	956.47	1.51	1.69E-39
CCDC80	5626.09	1.65	4.31E-39
ANKRD1	1359.09	1.54	3.27E-38
ID1	209.54	2.02	1.73E-34
SEMA5B	17.32	-4.91	4.10E-16
PTGDS	627.43	-2.80	7.90E-200
C3	2927.84	-2.07	1.92E-154
SERPINA1	1161.42	-2.12	1.88E-93
ITGB8	547.49	-2.28	6.77E-92
FXYD2	961.53	-2.15	1.30E-67
TNFSF10	961.02	-1.86	5.92E-66
CX3CL1	723.63	-1.65	1.07E-62
LTB	327.02	-2.56	3.27E-62
SOD2	1591.96	-1.79	2.37E-57
CYP1B1	491.70	-1.90	1.64E-54
FOS	221.60	-2.44	4.42E-53
TPCN1	866.23	-1.43	4.65E-45
PRODH	419.99	-1.61	1.25E-44
NPR1	927.00	-1.68	2.18E-44
SLCO4A1	258.34	-1.95	2.36E-42
C20orf75	1990.00	-1.25	1.39E-40
EDN2	207.77	-1.90	4.81E-40
PLAU	4235.28	-1.29	6.00E-38

Gene	Reads *	Log2 Fold Change	Adjusted P-value
CXCL2	290.97	-1.76	5.11E-36

Table 2Top 15 significantly changed KEGG pathways in response to TGF- β 1

Term	Pathway Size	Expected gene count	Observed gene count	P-value
Focal adhesion	201	23.22	50	5.89E-08
Cell adhesion molecules (CAMs)	133	15.37	36	5.09E-07
ECM-receptor interaction	84	9.70	26	1.30E-06
Complement and coagulation cascades	69	7.97	22	4.97E-06
Pathways in cancer	327	37.78	64	9.17E-06
Axon guidance	129	14.90	31	4.31E-05
Regulation of actin cytoskeleton	215	24.84	41	6.90E-04
Small cell lung cancer	84	9.70	20	1.10E-03
Leukocyte transendothelial migration	118	13.63	25	1.71E-03
TGF-beta signaling pathway	87	10.05	20	1.74E-03
Endocytosis	185	21.37	34	3.65E-03
p53 signaling pathway	68	7.86	16	3.82E-03
Chronic myeloid leukemia	75	8.66	17	4.39E-03
Hypertrophic cardiomyopathy (HCM)	85	9.82	18	7.29E-03
Propanoate metabolism	33	3.81	9	1.03E-02