

RESEARCH ARTICLE

Open Access



Analysis of the TGF β -induced program in primary airway epithelial cells shows essential role of NF- κ B/RelA signaling network in type II epithelial mesenchymal transition

Bing Tian^{1,2†}, Xueling Li^{3,4†}, Mridul Kalita^{1†}, Steven G. Widen^{2†}, Jun Yang¹, Suresh K. Bhavnani^{1,2,3}, Bryant Dang³, Andrzej Kudlicki^{2,3,4}, Mala Sinha^{2,4,5}, Fanping Kong^{2,4,5}, Thomas G. Wood^{2,3,4}, Bruce A. Luxon^{2,3,4,5} and Allan R. Brasier^{1,2,3*}

Abstract

Background: The airway epithelial cell plays a central role in coordinating the pulmonary response to injury and inflammation. Here, transforming growth factor- β (TGF β) activates gene expression programs to induce stem cell-like properties, inhibit expression of differentiated epithelial adhesion proteins and express mesenchymal contractile proteins. This process is known as epithelial mesenchymal transition (EMT); although much is known about the role of EMT in cellular metastasis in an oncogene-transformed cell, less is known about Type II EMT, that occurring in normal epithelial cells. In this study, we applied next generation sequencing (RNA-Seq) in primary human airway epithelial cells to understand the gene program controlling Type II EMT and how cytokine-induced inflammation modifies it.

Results: Generalized linear modeling was performed on a two-factor RNA-Seq experiment of 6 treatments of telomerase immortalized human small airway epithelial cells (3 replicates). Using a stringent cut-off, we identified 3,478 differentially expressed genes (DEGs) in response to EMT. Unbiased transcription factor enrichment analysis identified three clusters of EMT regulators, one including SMADs/TP63 and another NF- κ B/RelA. Surprisingly, we also observed 527 of the EMT DEGs were also regulated by the TNF-NF- κ B/RelA pathway. This Type II EMT program was compared to Type III EMT in TGF β stimulated A549 alveolar lung cancer cells, revealing significant functional differences. Moreover, we observe that Type II EMT modifies the outcome of the TNF program, reducing IFN signaling and enhancing integrin signaling. We confirmed experimentally that TGF β -induced the NF- κ B/RelA pathway by observing a 2-fold change in NF- κ B/RelA nuclear translocation. A small molecule IKK inhibitor blocked TGF β -induced core transcription factor (SNAIL1, ZEB1 and Twist1) and mesenchymal gene (FN1 and VIM) expression.

Conclusions: These data indicate that NF- κ B/RelA controls a SMAD-independent gene network whose regulation is required for initiation of Type II EMT. Type II EMT dramatically affects the induction and kinetics of TNF-dependent gene networks.

Keywords: Epithelial mesenchymal transition, Transforming growth factor β , Nuclear factor κ B, RNA-Seq, Generalized linear modeling, Transcription factor enrichment

* Correspondence: arbrasie@utmb.edu

†Equal contributors

¹Department of Internal Medicine, University of Texas Medical Branch (UTMB), Galveston, TX, USA

²Sealy Center for Molecular Medicine, UTMB, Galveston, TX, USA

Full list of author information is available at the end of the article

Background

The airway mucosa consists of highly polarized, differentiated epithelial cell types, whose primary role is to restrict fluid loss and limit inhaled particulate access to the internal milieu [1]. Exposure to aero-allergens, respiratory viruses or reactive oxygen stress, induces anchorage-dependent cell death (anoikis). During the process of anoikis, epithelial sloughing and disruption of the basement membrane releases sequestered extracellular matrix-associated epithelial growth factors, transforming growth factor (TGF), epidermal growth factor (EGF) and fibroblast growth factor (FGF). These epithelial growth factors activate gene expression programs in resident stem cell population (basal cells) to initiate epithelial repair and regeneration.

TGF β -stimulation of normal epithelial cells activates a de-differentiation program to promote cellular regeneration and extracellular matrix formation [1–3], referred to as Type II epithelial mesenchymal transition (EMT; [4]). Type II EMT induces loss of apical-basal polarity through cytoskeletal reorganization and dissolution of tight junctions. Concomitantly the cells express mesenchymal smooth muscle cell actin and intermediate filament vimentin (VIM) enhancing motility, and secrete collagen (Col1A), fibronectin (FN1), and matrix metalloproteinases (MMPs) promoting extracellular matrix deposition. Collectively this response promotes airway remodeling with expansion of the smooth muscle cell layer, repair of the epithelial surface and fibrosis [4, 5]. In this way, the airway epithelium plays a primary role in the resolution of injury or, if unresolved, the pathogenesis of chronic airway disease [6].

EMT is a dynamic and reversible epigenetic reprogramming event that plays a role in embryogenesis, organ homeostasis in response to acute injury, or cancer progression/metastasis, referred to as Types I, II, or III EMT, respectively [4]. Using Type III EMT as a model of cancer progression and metastases, the molecular signaling of how TGF β 1 initiates EMT has been studied extensively. Ligand induced activation of the transmembrane serine/threonine kinase TGF β receptor type II (TGF β RII) recruits and phosphorylates TGF β RI to signal through Smad-dependent, “canonical” and Smad-independent “noncanonical” pathways [7]. In the canonical pathway, phosphorylated Smad2/3 binds to Smad4 and the complex then translocates to the nucleus. The noncanonical signaling pathways involve downstream PI3K/Akt, Ras small GTPases, Wnt/ β -catenin, ERK, p38, and JNK. Collectively, the canonical and noncanonical pathways converge on a core set of transcription factors that function to initiate and maintain EMT: SNAIL (SNAI) 1/2, zinc finger E-box binding (ZEB) 1/2, Twist 1/2, Goosecoid and others [8].

The core EMT transcription factors mediate a coordinated series of gene activation/repression and chromatin

reprogramming events to shift cells to the mesenchymal phenotype. The activated Smad 2/3/4 trimer binds to Smad-binding elements in the regulatory regions of *SNAIL*, *JunB* and *c-Jun*, activating their expression by recruiting coactivators including the cAMP-response element binding protein (CBP)/p300 histone acetyltransferases [9]. SNAIL, in turn, represses *ECad* and zona occludin-1 genes by recruiting the polycomb complex, producing silencing histone modifications [10–12]. Smad signaling also increases expression of *ZEB1/2*, resulting in *MMP* and *collagen 1A* expression [8]. ZEB interacts with lysine-specific demethylase (LSD1), a protein involved in histone demethylation and chromatin reprogramming in EMT [13, 14]. Together these proteins coordinate both the repression of epithelial related genes and activation of mesenchymal genes.

Because of the temporal interplay of diverse signaling programs required to initiate and maintain EMT reprogramming, the EMT is highly modified by the state of cellular transformation and concomitant activation of extracellular signaling pathways. Oncogenic mutations in K-ras, activation of Wnt signaling, ROS stress and activation of insulin-like growth factor pathways that cross-talk with the TGF β pathway modify the expression of the EMT program [15]. As a result, the EMT program can be modulated by extracellular matrix interactions [16], and, of interest here, pro-inflammatory monocyte derived cytokines. TNF is a prototypical monokine [16, 17], whose actions trigger activation of p38 MAPK and JNK, essential components of the non-canonical TGF β signaling pathways [18, 19], and induce EMT in K-ras transformed epithelial cells through the actions of NF- κ B on the Twist EMT core transcription factor [16, 20]. However, the role of NF- κ B signaling in the EMT of normal epithelial cells is not known.

In this study we sought to examine the gene program of Type II EMT and to identify how this process was modulated by interaction with the innate signaling pathway. A well-established model of TGF β -induced EMT was applied to primary immortalized human small airway epithelial cells (hSAECs) to identify the gene expression networks responsible [5], and understand how activation of the innate response was modulated by EMT. Surprisingly, we observed that TGF β produced a gene expression program that was significantly enriched in NF- κ B-dependent genes identified by comparison to TNF dependent genes and to RelA enriched target genes in public ChIP-Seq data. Moreover, Type II EMT produces profound rewiring of the TNF gene program, skewing the pathway towards expression of integrin signaling to maintain the EMT state. We demonstrate that inhibiting NF- κ B/RelA via gene silencing or by inhibition of the IKK regulatory kinase blocked TGF β -induced EMT. These data indicate that NF- κ B/RelA gene

expression program is a major regulator of TGF β -induced Type II EMT.

Methods

hSAEC culture and EMT transformation

An immortalized human small airway epithelial cell (hSAEC) line was established by infecting primary hSAECs with human telomerase (hTERT) and cyclin dependent kinase (CDK)-4 retrovirus constructs [21]. The immortalized hSAECs were grown in SAGM small airway epithelial cell growth medium (Lonza, Walkersville, MD) in a humidified atmosphere of 5 % CO₂. For induction of EMT, hSAECs were TGF β stimulated for 15 days (10 ng/ml, PeproTech, Rocky Hill, NJ). The small molecule inhibitor of IKK, BMS345541 was purchased from Sigma Aldrich and used at 10 μ M [22].

Fluorescence microscopy

hSAECs were incubated in the absence or presence of TGF β (10 ng/mL) for 15 days, re-plated on glass cover slips pretreated with rat tail collagen (Roche Applied Sciences) and fixed with 4 % paraformaldehyde in PBS. Afterwards, the fixed cells were stained with Alexa Fluor[®] 568 phalloidin for cytoplasmic distribution of F-actin (shown in red color) and also counterstained with 4', 6-diamidino-2-phenylindole (DAPI) for nuclear staining (shown in blue color). The cells were visualized with a Nikon fluorescence confocal microscope at a magnification of 63 \times [5, 23].

Quantitative real time PCR (Q-RT-PCR)

Reverse transcription was performed on 1 μ g of total RNA with random primers, utilizing SuperScript III first-strand synthesis system for RT-PCR (Life technologies, Invitrogen) under conditions recommended by the

manufacturer. Equivalent amounts of RNA were assayed by quantitative (Q) Q-PCR. The PCR reaction consisted of SYBR Green[®] PCR Master Mix (Bio-Rad), template cDNA and assay primers (Table 1) in a total reaction volume of 20 μ l. Thermal cycling (50 $^{\circ}$ C, 2 min; 95 $^{\circ}$ C, 10 min; and 40 cycles at 95 $^{\circ}$ C, 15 S; 60 $^{\circ}$ C, 1 min) was performed using a MyiQ Single Color Real-Time PCR Detection System (Bio-Rad). Threshold cycle numbers (Ct) were defined as fluorescence values, generated by SYBR green binding to double stranded PCR products, exceeding baseline. Relative transcript levels were quantified as a comparison of measured Ct values for each reaction [24], normalized using cyclophilin as an internal control.

RNA extraction and qualification

Total cellular RNA was extracted using either RNAqueous[™] phenol-free total RNA isolation kits (Life Technologies, CA) or Quick-RNA MiniPrep kits (ZYMO Research) according to the manufacturer's recommendations. RNA was quantified spectrophotometrically using a NanoDrop ND-1000 (NanoDrop Technologies, DE). Quality of the purified RNA was assessed by visualization of 18S and 28S RNA bands using an Agilent BioAnalyzer 2100 (Agilent Technologies, CA). The resulting electropherograms were used in the calculation of the 28S/18S ratio and the RNA Integrity Number.

Library construction and sequencing

Poly-A⁺ RNA was selected from total RNA (1 μ g) using oligo dT-attached magnetic beads. Bound RNA was fragmented by incubation at 94 $^{\circ}$ C for eight (8) minutes in 19.5 μ l of fragmentation buffer (Illumina). First and second strand synthesis, adapter ligation and amplification of the library were performed using the Illumina TruSeq

Table 1 PCR Primers for genes in Q-RT-PCR analysis

Primer Set	Sequence (5'- 3')	
	Forward	Reverse
hCol1A	CCAGAAGAAGTGGTACATCAGCA	CGCCATACTCGAACTGGAATC
hFN1	CGGTGGCTGTCAGTCAAAG	AAACCTCGGCTTCTCCATAA
hIL-6	CTGGATTCAATGAGGAGACTTGC	TCAAATCTGTTCTGGAGTACTCTAGG
hSNAIL1	GCGCTCTTTCCTCGTCAGG	GGGCTGCTGGAAGGTAAACTCT
hTwist1	TCTCGGTCTGGAGGATGGA	CAATGACATCTAGGTCTCCG
hVIM	GCTCAATGTTAAGATGGCCCTT	TGGAAGAGGCAGAGAAATCTCG
hZEB1	GATGATGAATGCGAGTCAGATGC	ACAGCAGTGTCTTGTGTGTGT
hPPIA	CCCACCGTGTCTTCGACATT	GGACCCGTATGCTTTAGGATGA
mIL-6	TAGTCCTTCTACCCCAATTTC	TTGGTCCTTAGCCACTCCTTC
mSNAIL1	CACACGCTGCCTTGTGTCT	GGTCAGCAAAGCACGGTT
mZEB1	ACCGCCGTCATTTATCCTGAG	CATCTGGTGTCCGTTTTCATCA
mPPIA	GAGCTGTTGCAGACAAAGTTC	CCCTGGCACATGAATCCTGG

RNA Sample Preparation kit under conditions prescribed by the manufacturer (Illumina). Samples were tracked using “index tags” incorporated into the adapters. Library quality was evaluated using an Agilent DNA-1000 chip on an Agilent 2100 Bioanalyzer. Quantification of library DNA templates was performed using qPCR and a known-size reference standard. Cluster formation of the library DNA templates was performed using the TruSeq PE Cluster Kit v3 (Illumina) and the Illumina cBot workstation using conditions recommended by the manufacturer. Template input was adjusted to obtain a cluster density of 700–900 K/mm². Paired end 50 base sequencing by synthesis was performed using TruSeq SBS kit v3 (Illumina) on an Illumina HiSeq 1000 using protocols defined by the manufacturer.

GLM modeling and pathway analysis

As shown in Table 2, the experimental design has six sample groups and requires a two factor analysis for time of TNF stimulation (“Time”) and presence of TGF β stimulation (“Transformed”). For this purpose we employed the Generalized Linear Modeling (GLM) capabilities in the R Bioconductor package edgeR [25] to perform the data modeling phase of the analysis within our data analysis pipeline shown in Fig. 2b. EdgeR was developed to analyze differential count data arising from designed experiments such as RNA-Seq with single and multiple experimental factors and small numbers of replicates (here $n = 3$). It includes important statistical methods such as scaled normalization using trimmed means, the negative binomial distribution to describe read count variability, estimates of gene specific dispersion parameters by conditional maximum likelihood using Empirical Bayes methods, exact tests for differential gene expression analysis (DEGA) of one factor experimental designs and GLM for DEGA on multiple factor designs.

The raw NGS analysis passed eleven separate quality analyses by FastQC [26]. Reads were aligned using TopHat2 [27], a fast splice junction mapper that aligns RNA-Seq reads using the ultra-high-throughput short read aligner Bowtie2 [28], using the Burrows-Wheeler index method. The alignment across all 18 samples

produced a mean Overall Read Alignment Rate of $96.4 \pm 0.4 \%$ and a mean Concordant Pair Alignment Rate of $88.0 \pm 0.2 \%$, confirming that the read alignment was excellent and that data was of high quality. (The results of QA/QC of RNA Seq are provided at Additional file 1).

As shown in Fig. 2b, edgeR used the resulting RNA-Seq read count data to build the counts table in edgeR to begin DGEA. Our experimental design has six groups of cell types (Table 2) each comprised of three samples, thereby giving a total of 18 sample columns with 23,710 rows of read counts, each row representing a unique gene in the count table.

Data filtering by removing unresponsive (i.e., uninformative) genes, was performed to improve the power of the study by reducing effects of multiple testing correction. We use this step to eliminate seemingly unresponsive genes (rows) from the count table using our rule that to be considered responsive, at least one of the six groups has to have a significant number of counts in at least two of its three samples. This filtering step culled out 7,346 rows resulting in a final read count table having 16,364 genes with a mean library size of $4.54E^7$ across the 18 samples. This reduced read count table resulted in a mean scaling factor of 1.002 (range: 0.7985 – 1.0990).

Before creating a model design, we identified the model factors (Transformed, Time) and levels to edgeR as [N.0 h, N.1 h, N.12 h, Y.0 h, Y.1 h, Y.12 h] where N = no TGF β transformation, Y = transformation with TGF β and hr = time post TNF α treatment in hours. From these factors and levels we can identify the important contrasts that we would like to explore in our DEGA as shown in Table 3. We then designed the statistical model that was most appropriate for guiding the analysis in edgeR. This approach allowed us to optimize the information content while minimizing type I and type II errors, so we can then select those genes most suitable for pathways analysis.

The edgeR GLMs are non-linear models requiring iterative fitting and dispersion estimation to allow it to account for variations in gene abundance between RNA samples. EdgeR uses the Cox-Reid profile-adjusted

Table 2 Abbreviations and descriptions of the sample groups for differential gene expression (DGE) analysis

Code	Description
CE	Control (C) untransformed Epithelial cells (E)
CM	Control (C) TGF β transformed Mesenchymal cells (M)
T1E	TNF α treated (T) E cells 1 h post treatment
T1M	TNF α treated (T) M cells 1 h post treatment
T12E	TNF α treated (T) E cells 12 h post treatment
T12M	TNF α treated (T) M cells 12 h post treatment

Table 3 Compared sample group pairs (contrasts) in the DGE analysis

Contrasts
CM–CE
T1E–CE
T12E–CE
T12E–T1E
T1M–T1E
T12M–T12E
T1M–CM
T12M–CM
T12M–T1M

likelihood method in estimating dispersions and the read count data is modeled using the negative binomial distribution because it accounts for overdispersion, which the Poisson model does not. This is important in RNA-Seq experiments because the observed variance (i.e., dispersion) in the read count data is often greater than that predicted from statistical theory. The tagwise dispersion was estimated in a gene-wise fashion in three steps using empirical Bayes shrinkage via a weighted likelihood method. Our estimated common dispersion was 0.01548079 which suggests that our sequencing quality is very good.

We next performed the GLM fit using the statistical design then performed the GLM Likelihood Ratio Test (LRT) for each of the contrasts in Table 3. Using each of the LRT outputs, we then performed Bonferroni FWER [29] and Benjamini and Hochberg FDR [30] multiple testing corrections on the DGE for each contrast using the edgeR topTags method. For further analysis such as pathway analysis, gene selection stringency is a key issue so we used a 2× absolute fold change cutoff before p-value filtering. This was a very strong response field and was amenable to a very high level of stringency so we focused on the Bonferroni corrected results with a FWER p-value cutoff of 0.00001 to insure that genes chosen for further analysis were reliably responsive to our experimental conditions and to keep the number of genes in the downstream analytical pool reasonable. Differentially expressed genes are shown in Additional files 2, 3, 4, 5, 6, 7, 8, and 9.

Venn diagrams were created to examine the intersections and relative complements of groups of contrasts (treatment pairs compared) that were deemed biologically interesting using locally developed tools. Subsets of genes identified as interesting were explored using QIAGEN's Ingenuity® Pathways Analysis suite (IPA®, QIAGEN Redwood City, www.qiagen.com/ingenuity) for pathways, networks, and functional analyses.

Transcription factor enrichment analysis

ChIP-X data were downloaded from the website <http://amp.pharm.mssm.edu/lib/chea.jsp> in April 2014. The ChIP-X data consist of 148 transcription factors (TFs) and their respective target genes (TGs) derived from ChIP-chip, ChIP-Seq, ChIP-PET or DamID experiments in 237 publications [31]. Since TFs in different cell lines and treatments may result in different TF binding data, we separated experiments on the same TF, resulting in 345 combinations of TF and experiments performed on human, mouse and rat cells. Among the 345 TF binding experiments, 148 were performed on human. Here we refer to the ChIP-X data as the ChEA ChIP-X data set.

We also downloaded the uniform histone and TF binding peaks from the Encyclopedia of DNA Elements (ENCODE)

Consortium website (http://ftp.ebi.ac.uk/pub/databases/ensembl/encode/integration_data_jan2011/byDataType/signal/jan2011/bigwig/) to serve as a second data set. As described above, we separated binding peaks of the same TF from different experiments, which results in 1,169 distinct ChIP-Seq experiments. The target genes corresponding to the TF binding peaks were identified according to PAVIS criteria by using our own scripts [32]. Here the coordinates of the peaks were mapped based on the hg19 refFlat.txt file downloaded from UCSC on Feb. 2012.

Differentially expressed genes (DEGs) between epithelial cells and mesenchymal cells were identified with FPKM fold change ratio ≥ 2 or ≤ -2 and q-value ≤ 0.01 from the RNA-Seq data. 3,487 differentially expressed genes were identified among the 23,284 genome-wide genes considered. Among the DEGs, 2010 genes are downregulated and 1477 are upregulated.

To determine if the transcriptional program regulated by a specific transcription factor from the above 148 transcription factors was activated in the EMT process or not, the hypergeometric probability distribution was used as the background distribution. The enrichment for each transcription factor regulated target gene (TG) was calculated by $p\text{-value} = 1 - \text{hygecdf}(k-1, N, K, n)$ using the Matlab build-in function `hygecdf`. The number of overlapping genes between the DEGs and the target genes for the specific transcription factor is k ; N is the number of genome-wide genes: 23,284; K , the number of the TGs of the TF in the ChIP-data; n , the number of DEGs. For multiple testing corrections, Bonferroni FWER was then performed by multiplying the p-value with the number of TF binding experiments, i.e., 345 TF and experimental combinations used in the analysis based on the ChEA ChIP-X data and 1,169 that based on ENCODE ChIP-Seq data as the corrected p-value. Enrichment fold ratio was calculated by $k/n/(K/N)$. If the corrected p-value by Bonferroni correction was smaller than 0.01 and fold ratio is greater than 1.5, the corresponding transcription factor was considered significantly enriched. For the TF enrichment analysis on ENCODE ChIP-Seq data, a Benjamini-Hochberg correction was also performed to estimate the false discovery rate (FDR) and q-value.

Hierarchical clustering analysis

TF and target genes were subjected to hierarchical clustering analysis using the Euclidean distance measure and the Ward. D2 linkage function in Spotfire (TIBCO). Variance, Skewness, and Kurtosis of dendrogram results were compared to 1000 random permutations of the data.

Comparison of type II and type III EMT programs

We compared DEGs in Type II EMT with a publicly available dataset from a model of Type III EMT (GSE177708, NCBI Gene Expression Omnibus). Affymetrix microarrays

were used to measure gene expression in human adenocarcinoma cell line A549 after TGF β induction of EMT. Genes in the two datasets were matched by Entrez ID, and only genes that were common to both datasets, and detected to be expressed in the microarray, were kept for further analysis.

Results

Induction of the EMT program

We have previously validated a model of Type II EMT using a continuously replicating line of human small airway epithelial cells (hSAECs) immortalized with human telomerase (hTERT) and CDK4 expression [21]. These cells show a stable epithelial morphology and differentiated cytokeratin isoforms after over 100 population doublings, express the stem cell marker p63 and high levels of p16INK4a, and have an intact p53 checkpoint pathway [21]. In the absence of TGF β stimulation, hSAECs assumed a normal cuboidal morphology with perinuclear cytoplasmic distribution of F-actin, detected by fluorescence microscopy using confocal microscopy after staining with Alexa568-conjugated phalloidin (Fig. 1a). In response to chronic TGF β stimulation, the cells acquired an elongated shape with markedly induced F-actin staining. This morphological change of enhanced front-rear polarity and actin rearrangement into cytosolic

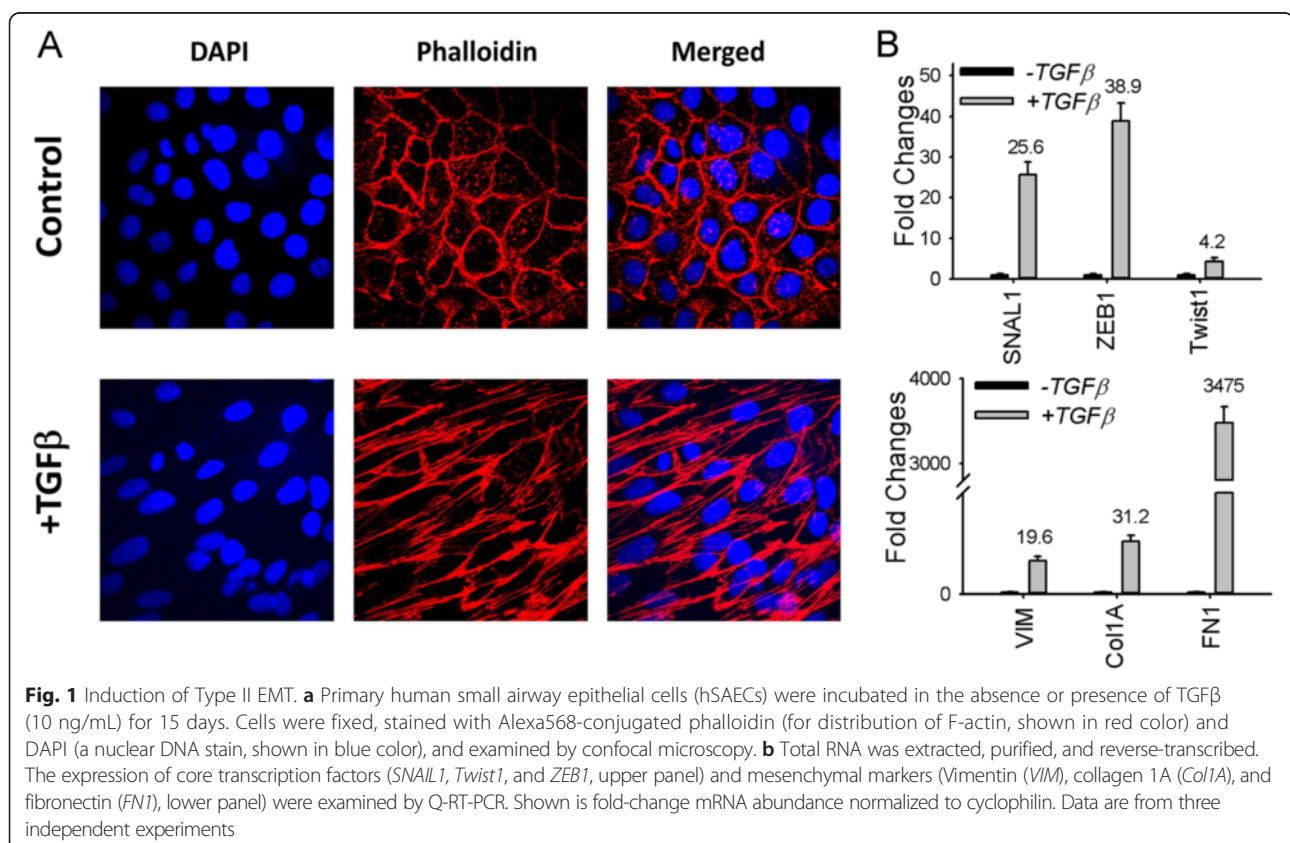
stress fibers are characteristic morphological changes of EMT [5, 33].

To demonstrate the induction of the Type II EMT gene program, the steady state expression of mesenchymal genes and EMT-associated transcription factors were assessed by Q-RT-PCR. TGF β -treatment induced the expression of *FN1*, *Col1A* and *VIM* by 3,475-, 31- and 20-fold, respectively (Fig. 1b). Similarly, the EMT core transcription factors *SNAIL1*, *ZEB1* and *Twist1* were induced by 26-, 39- and 4-fold, respectively (Fig. 1b). Together these data suggest that TGF β induces morphological and gene signatures of stable Type II EMT in hSAECs.

Analysis of the EMT gene program

To understand the core EMT program and how TNF signaling modulates its expression, hSAECs were subjected to TGF β induced reprogramming, and stimulated in the absence or presence of TNF for 1 or 12 h. Each experimental condition was reproduced in triplicate, representing 18 samples subjected to RNA-Seq analysis (Fig. 2a). The schema used in RNA sequence analysis is described in Fig. 2b.

Multidimensional scaling (MDS) was used as an unsupervised approach to visualize the level of similarity of the 18 RNA samples (Fig. 2c). MDS is a method similar



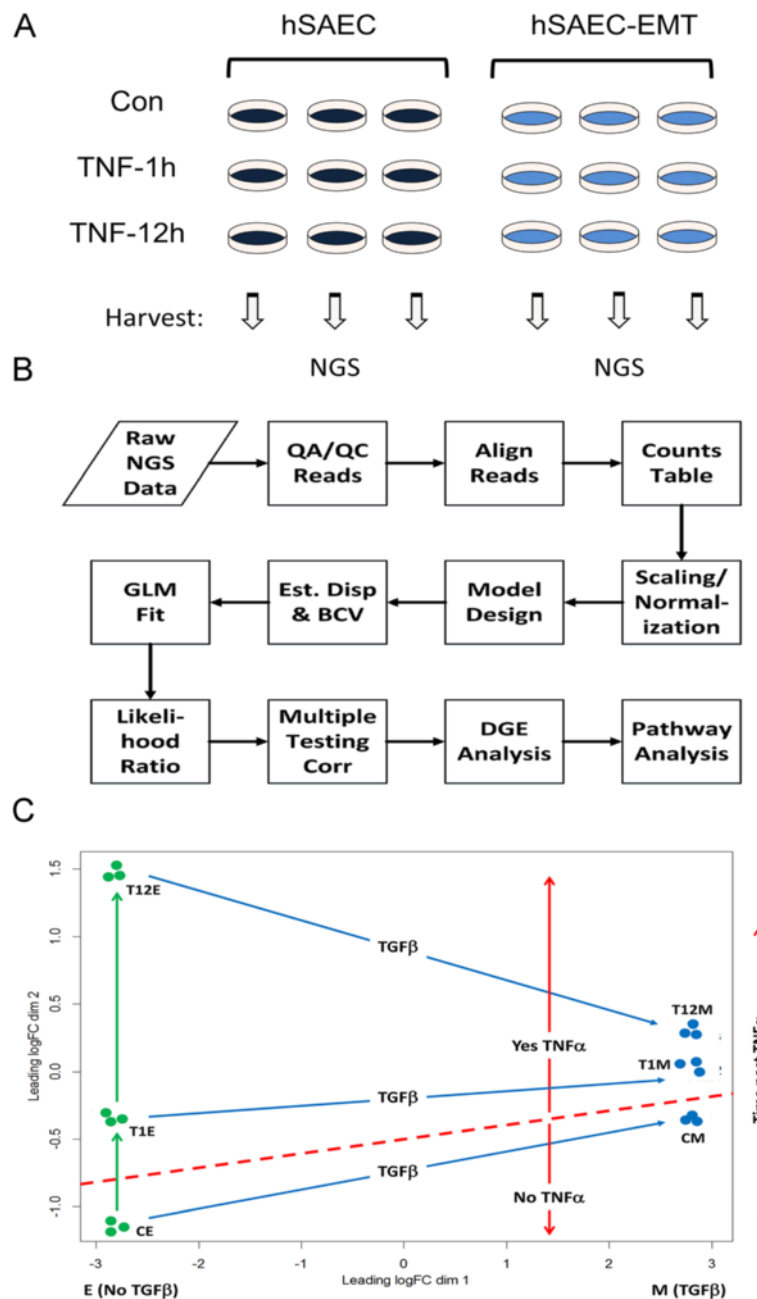


Fig. 2 Experimental perturbation and RNA-Seq analysis. **a** Experimental design. Triplicate biological replicates were subjected to no treatment (hSAEC) or TGFβ treatment (hSAEC-EMT). Cells were stimulated with TNFα for 0, 1 or 12 h prior to RNA extraction. **b** Flow diagram for RNA-Seq GLM data analysis. Abbreviations: NGS, next generation sequencing; QA/QC, quality assurance/quality control; GLM, generalized linear model; DGE, differential gene expression. **c** Multidimensional unsupervised clustering analysis of RNA-Seq data sets. Shown is the similarity of each replicate based on TGF treatment or time post TNF and their clustering by group. Abbreviations: CE, control (unstimulated) epithelial cells; CM, control mesenchymal cells

to principal component analysis that enables high level representation of RNA Seq data based on similarities in expression of 500 genes with the largest fold change difference in the data set. In this analysis, distances were calculated as leading log-fold-changes between each sample pair. From Fig. 2c it is readily apparent that

MDS differentiates the RNA seq expression patterns of the mesenchymal state from that of the epithelial cells in the first dimension and that of TNFα in the second dimension. Although TNF induces very strong effect on the epithelial cells in the second dimension, the relative response of the mesenchymal cells is highly compressed

in the second dimension. Finally, the differences in fold change between replicates in the same group are very small relative to the inter-group responses. These analyses indicated that we had a robust and reproducible dataset and that chronic TGF β treatment partially appeared to mimic acute stimulation of epithelial cells with TNF.

Generalized linear modeling was used to identify genes that were differentially expressed in the data set; contrasts were used to compare the differences between treatment groups. The results of generalized linear modeling (GLM) are provided in Additional files 2–10. We first focused on the comparison of the mesenchymal vs epithelial states. This comparison identified 3,487 differentially expressed genes (DEGs) that met a fold change cut-off of >2 with a highly significant, Bonferroni corrected pvalue ($p < 0.00001$). This group of DEGs was subjected to Ingenuity Pathway Analysis (IPA), where canonical pathways were ranked by the enrichment of genes in each process. Here we noted the rank ordered appearance of “cell death and survival”, “organismal injury and abnormalities”, and “cellular development” were influenced by the TGF β -induced EMT (c.f. Fig. 3a, 3b).

Analysis of canonical pathways further showed that NF- κ B/RelA-signaling pathway was significantly enriched in a network with SP1 and STATs (Fig. 3c), factors known to modulate RelA-dependent signaling [34, 35].

Transcription factor enrichment of the EMT

An unbiased transcription factor (TF) enrichment analysis was performed on the DEGs. In this analysis, 345 distinct TF experiments of ChIP-chip, ChIP-Seq, ChIP-PET or DamID were tested for TF enrichment using hypergeometric probability distribution [31]. Here, we only consider the enriched TFs from 148 DNA binding experiments performed on human cells. Of these, 47 distinct TF experiment datasets were significantly enriched for the 3,487 DEGs in EMT. The enriched TFs are shown in Table 4, ranked by enrichment fold ratio and associated with the number of upregulated genes. Many of these TFs are linked to EMT in the literature. Of these, SMADs were among the top highly enriched TFs, consistent with their known role in mediating canonical TGF β signaling [16]. Also included in the top list was tripartite motif containing 28 (TRIM28), an epigenetic modulator regulating histone H3 modifications on E-cadherin and N-cadherin promoters [36]. SOX2 promotes tumor metastasis by stimulating EMT via regulation of the WNT/ β -catenin signal network [37]. ESR1 and ESR2, reported to suppress EMT [38, 39], were also significantly enriched. The CLOCK TF binds the E-box of *PER2* and regulates this and other circadian rhythm genes, which are reported to be involved in EMT [40].

The tumor suppressor TP53 is known to regulate EMT [41, 42]. We observed that p53-related TP63, IRF1, TFAP2C, CTNNB1, AHR1, ELF5, FOXM1 and

FOXO3 were also exclusively enriched TFs for downregulated DEGs (Table 5). Among these, TP63 attenuates EMT in prostate cells [43] and alternatively spliced isoform Δ Np63 α inhibits EMT in human bladder cancer cells [44]. IRF1 is a tumor suppressor modulated by miR-23, promoting TGF-beta-induced EMT in lung cancer [45]. TFAP2C governs the luminal epithelial phenotype in mammary development and carcinogenesis [46]. AHR1 was reported to inhibit TGF β -induced EMT [47]; ELF5 inhibits Type I EMT in mammary gland development and Type III EMT in breast cancer by transcriptionally repressing SNAIL2 [48]. Constitutive activation of CTNNB1 stabilized mesenchymal phenotypes of epithelial cells [49]. Of relevance to the hypothesis developed in this work, we noted that NF- κ B/RelA was among the top significantly enriched TFs, with ChIP-Seq binding sites for 152 of the 1477 upregulated genes in the EMT data set, indicating that NF- κ B/RelA dependent genes may play an important role in Type II EMT.

The depleted TFs based on the ChEA ChIP-X data set include ELF1, GABP, KDM6A, EST1, ETS1 (Tables 6, 7), suggesting that these TFs may be involved in the reversal of EMT, termed MET. GABP has a predicted binding site on miR200c, an important effector of EMT in immortalized mammary epithelial cells, MCF12A [41]. KDM6A was reported to regulate the expression of TFs critical for stem cell differentiation. EMT is a process of dedifferentiation and gain of stemness, which may explain the depletion of this TF in Type II EMT. EST1 and ETS1 have not yet been reported to be involved in MET. Although our TF enrichment results may be affected by limited experiments in the public databases of human epithelial cells, this analysis has generated valuable hypotheses on the TFs involved in EMT, most of which are supported by the literature.

Further comparison of the TF enrichment analysis indicated that the 30 enriched TFs for upregulated DEGs significantly overlapped with the 40 TFs associated with downregulated DEGs (p -value = $4.3e-13$), with 18 TFs in common (Tables 4 and 5). A similar finding was observed with depleted TFs in the EMT DEGs (Tables 6 and 7), suggesting that many transcription factors are bimodal, i.e., they both inhibit or activate the target genes depending on the chromatin environment.

Topographical map of EMT regulated transcription factors and target genes

We next sought to map the relationships between TFs and target genes controlling EMT. In this analysis, hierarchical clustering identified three significant TF clusters and two DEG clusters (Fig. 4). The gene clusters correspond to genes that were either up- or downregulated by EMT. Conversely, we noted that each TF is associated with both up- and down-regulated target genes, indicating

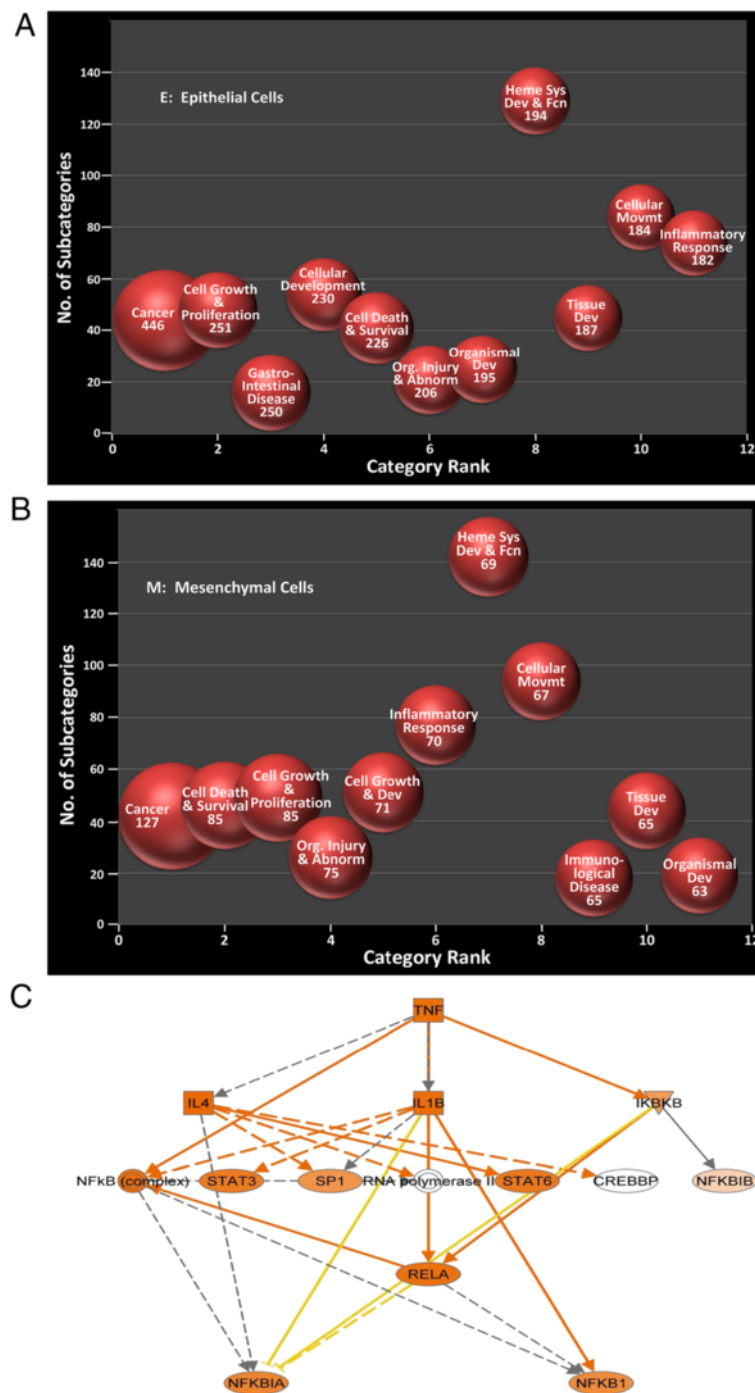


Fig. 3 Functional analysis of differentially expressed genes in EMT. **a** Ingenuity pathway analysis (IPA) is shown for genes expressed in control hSAECs. X axis, category rank; Y axis, number of subcategories in each pathway. **b** IPA for genes in hSAEC-EMT. Note the cell death and survival and organismal injury and abnormality become the 2nd and 4th ranked pathway. **c** Enriched TNF signaling pathway identified by IPA in DEG data set. Intensity of red hue is proportional to upregulation by EMT

bimodal behavior consistent with the above TF enrichment analysis. The TF cluster A contains TP63 and Smad2/3. Both Smad 2 and 3 regulate relatively fewer DEGs, some of which are coregulated by EGR and SOX2.

TF cluster B contains RelA, BTB and CNC homology 1 (BACH1), globin transcription factor (GATA1/2), and MYC proteins; TF cluster B regulates a distinct group of DEGs from those regulated by cluster A, but also a

Table 4 Significantly enriched human transcription factors for upregulated genes in the EMT process. The enriched transcription factors (TFs) were inferred by comparing ChIP-Seq or ChIP-chip identified target genes (TGs) of each TF with the upregulated DEGs in the EMT process. The enriched TFs were sorted according to the enriched fold ratios

Human TFs	PMID	# total TGs	# DETGs	Fold ratio	Corrected p-value	Cell lines
TRIM28	21343339	85	21	3.9	1.8E-05	HEK293
ESR1	21235772	213	36	2.7	4.8E-04	MCF-7
CLOCK	20551151	398	58	2.3	2.3E-05	293 T
ESR2	21235772	392	57	2.3	9.1E-07	MCF-7
SMAD3	21741376	1074	144	2.1	1.4E-06	HESC
SOX2	20726797	2060	275	2.1	<3.45E-14	SW620
RELA	24523406	1147	152	2.1	<3.45E-14	A549
BACH1	22875853	1183	150	2	<3.45E-14	HELA-AND-SCP4
GATA1	19941826	700	84	1.9	3.8E-14	K562
TP53	18474530	734	88	1.9	4.4E-06	U2OS
GATA2	21666600	634	75	1.9	2.1E-06	HMVEC
PAX3-FKHR	20663909	975	115	1.9	4.8E-05	RHABDOMYOSARCOMA
SMAD4	21741376	2032	237	1.8	2.4E-08	HESC
DROSHA	22980978	441	51	1.8	<3.45E-14	HELA
ZNF217	24962896	1400	160	1.8	8.7E-03	MCF7
NFE2L2	22581777	508	58	1.8	4.7E-11	LYMPHOBLASTOID
EOMES	21245162	861	98	1.8	3.7E-03	HESC
SMAD2	18955504	1643	184	1.8	4.1E-06	HaCaT
SMAD3	18955504	1643	184	1.8	3.8E-12	HaCaT
NR3C1	21868756	1023	111	1.7	3.8E-12	MCF10A
DNAJC2	21179169	803	87	1.7	6.1E-06	NT2
EGR1	20690147	5216	549	1.7	2.4E-04	ERYTHROLEUKEMIA
SOX2	21211035	2830	288	1.6	<3.45E-14	LN229_GBM
GATA2	19941826	1727	175	1.6	<3.45E-14	K562
MYC	19915707	1793	177	1.6	8.5E-08	AK7
AR	19668381	2958	291	1.6	5.0E-07	PC3
ATF3	23680149	1995	193	1.5	8.8E-13	GBM1-GSC
SMAD4	21799915	2302	220	1.5	4.3E-07	A2780
GATA1	21571218	2388	228	1.5	7.2E-08	MEGAKARYOCYTES

PMID pubmed article number; # total TGs number of ChIP-Seq or ChIP-chip identified target genes (TGs) of each TF; #DETs number of differentially expressed genes

bimodal fashion. Cluster C is the largest group of TFs containing estrogen receptor (ESR1/2), forkhead box 03 (FOXO3), TRIM28 and others, and is associated with the fewest up- or downregulated DEGs.

Type II EMT program controls biological functions distinct from than that of Type III EMT

Because Type II EMT occurs in primary epithelial cells without background oncogenic K-ras mutations characteristic of epithelial tumors [50], we hypothesized that Type II EMT induces distinct gene expression profiles. Differences in gene regulation between Type II EMT and Type III EMT may help identify pathways that are

specifically involved in fibrosis (Type II EMT) vs invasion and metastases (Type III EMT). We therefore compared the TGF β -induced gene expression patterns in primary hSAEC with that of an established human epithelial alveolar carcinoma cell, A549. A549 cells are K-Ras activated, keratin positive epithelial cells with features of the lower airways [51, 52]. For this purpose, we compared our Type II model dataset with a publicly available Type III model dataset (GSE177708, NCBI Gene Expression Omnibus). Affymetrix microarrays were used to measure gene expression in Human adenocarcinoma cell line A549 after TGF β induction of EMT. Genes in the two datasets were matched by Entrez ID,

Table 5 Significantly enriched human TFs for the downregulated genes in the EMT process. The enriched transcription factors (TFs) were inferred by comparing ChIP-Seq or ChIP-chip identified target genes (TGs) of each TF with the downregulated DEGs in the EMT process. The enriched TFs were sorted according to the enriched fold ratios

Human TFs	PMID	# total TGs	# DETGs	Fold ratio	Corrected <i>p</i> -value	Cell lines
TP63	17297297	34	14	3.9	1.1E-04	HaCaT
TRIM28	21343339	85	22	3	7.4E-04	HEK293
ESR2	21235772	392	98	2.9	<3.45E-14	MCF-7
CTNNB1	24651522	138	30	2.5	6.5E-04	LGR5+ INTESTINAL STEM CELL
SOX2	20726797	2060	433	2.4	<<3.45E-141.0E-16	SW620
IRF1	21803131	326	67	2.4	5.9E-09	MONOCYTES
TFAP2C	20629094	1125	229	2.4	<3.45E-14 < 3.45E-14 < 3.45E-14	MCF7
ESR1	21235772	213	43	2.3	4.1E-05	MCF-7
ZNF217	24962896	1400	271	2.2	<3.45E-14	MCF7
FOXM1	23109430	253	47	2.2	1.4E-04	U2OS
NR1H3	23393188	574	104	2.1	8.6E-11	ATHEROSCLEROTIC-FOAM
CDX2	20551321	383	69	2.1	1.2E-06	CACO-2
BACH1	22875853	1183	211	2.1	<3.45E-14	HELA-AND-SCP4
AHR	22903824	621	109	2	1.9E-10	MCF7
FOXO3	23340844	650	114	2	5.9E-11	DLD1
AR	21909140	258	45	2	1.5E-03	LNCAP PROSTATE CANCER CELL LINES
TP53	16413492	276	48	2	7.8E-04	HCT116
SMAD2	18955504	1643	285	2	<3.45E-14	HaCaT
SMAD3	18955504	1643	285	2	<3.45E-14	HaCaT
TP53	18474530	734	122	1.9	3.9E-10	U2OS
RELA	24523406	1147	187	1.9	<3.45E-14	A549
TP63	23658742	3355	541	1.9	<3.45E-14	EP156T
ARNT	22903824	912	147	1.9	2.2E-11	MCF7
SMAD4	19686287	348	56	1.9	1.4E-03	HaCaT
TCF4	18268006	392	63	1.9	3.9E-04	LS174T
CLOCK	20551151	398	62	1.8	1.3E-03	293 T
PAX3-FKHR	20663909	975	144	1.7	3.1E-08	RHABDOMYOSARCOMA
TP63	22573176	3867	570	1.7	<3.45E-14	HFKS
GATA1	19941826	700	103	1.7	1.9E-05	K562
SMAD3	21741376	1074	158	1.7	4.8E-09	HESC
CTNNB1	20460455	900	130	1.7	1.1E-06	HCT116
FOXA2	19822575	2791	402	1.7	<3.45E-14	HepG2
SMAD4	21741376	2032	290	1.7	<3.45E-14	HESC
ELF5	23300383	920	130	1.6	4.2E-06	T47D
GATA2	19941826	1727	241	1.6	5.2E-12	K562
TRIM28	17542650	680	94	1.6	1.2E-03	NTERA2
GATA2	21666600	634	87	1.6	3.6E-03	HMVEC
AR	19668381	2958	398	1.6	<3.45E-14	PC3
HNF4A	19761587	816	108	1.5	1.7E-03	HUMAN INTESTINAL CELL LINE CACO-2
AR	22383394	1690	221	1.5	4.3E-08	PROSTATE_CANCER

PMID pubmed article number; # total TGs number of ChIP-Seq or ChIP-chip identified target genes (TGs) of each TF; #DEGs number of differentially expressed genes

Table 6 Significantly depleted human TFs for the upregulated genes in the EMT process

Human TFs	PMID	# total TGs	# DETGs	Fold ratio	Corrected <i>p</i> -value	Cell lines
ELF1	17652178	99	0	Inf	<3.45E-14	JURKAT
GABP	17652178	655	9	4.6	2.2E-08	JURKAT
EST1	17652178	636	9	4.5	6.4E-08	JURKAT
ETS1	20019798	1446	28	3.3	2.8E-14	JURKAT
KDM6A	18722178	410	9	2.9	7.4E-03	U937_AND_SAOS2
FOXP3	21729870	1300	43	1.9	6.1E-05	TREG
VDR	23849224	2029	69	1.9	9.1E-08	CD4+
GABP	19822575	2430	92	1.7	7.2E-07	HepG2
HOXC9	25013753	1858	72	1.6	1.7E-04	NEUROBLASTOMA BE2-C

The depleted transcription factors (TFs) were inferred by comparing ChIP-Seq or ChIP-chip identified target genes (TGs) of each TF with the upregulated DEGs in the EMT process. Depleted TFs were sorted according to the depleted fold ratios

PMID pubmed article number; # total TGs number of ChIP-Seq or ChIP-chip identified target genes (TGs) of each TF; #DETGs number of differentially expressed genes

and only genes that were common to both sets, and detected to be expressed in the microarray, were kept for analysis.

We identified 137 genes upregulated in Type II- but not in Type III-EMT, and 124 genes upregulated in Type III- but not in Type II EMT. Fig. 5 (top) illustrates the top 10 canonical pathways for genes uniquely induced in Type II EMT. In this plot, both the *p* value and ratio of enrichment are presented for each pathway. Here, lipid metabolism and thrombin signaling pathways were the most statistically significant. By contrast, Fig. 5 (bottom) displays the top 10 canonical pathways for genes uniquely induced in Type III EMT. Plasma membrane estrogen receptor signaling, ErbB receptor signaling and glycipan pathways were the most highly significant, consistent with the effect of Type III EMT on cellular motility, invasion and metastasis. Together these data suggest that the induction of Type II EMT produces a cellular biology state distinct from that seen in Type III EMT.

Modulation of the TNF response by Type II EMT

NF- κ B is an inducible transcription factor that is regulated by a rapid release from cytoplasmic I κ B α stores, followed by a slower release from 105 kDa NF κ B2

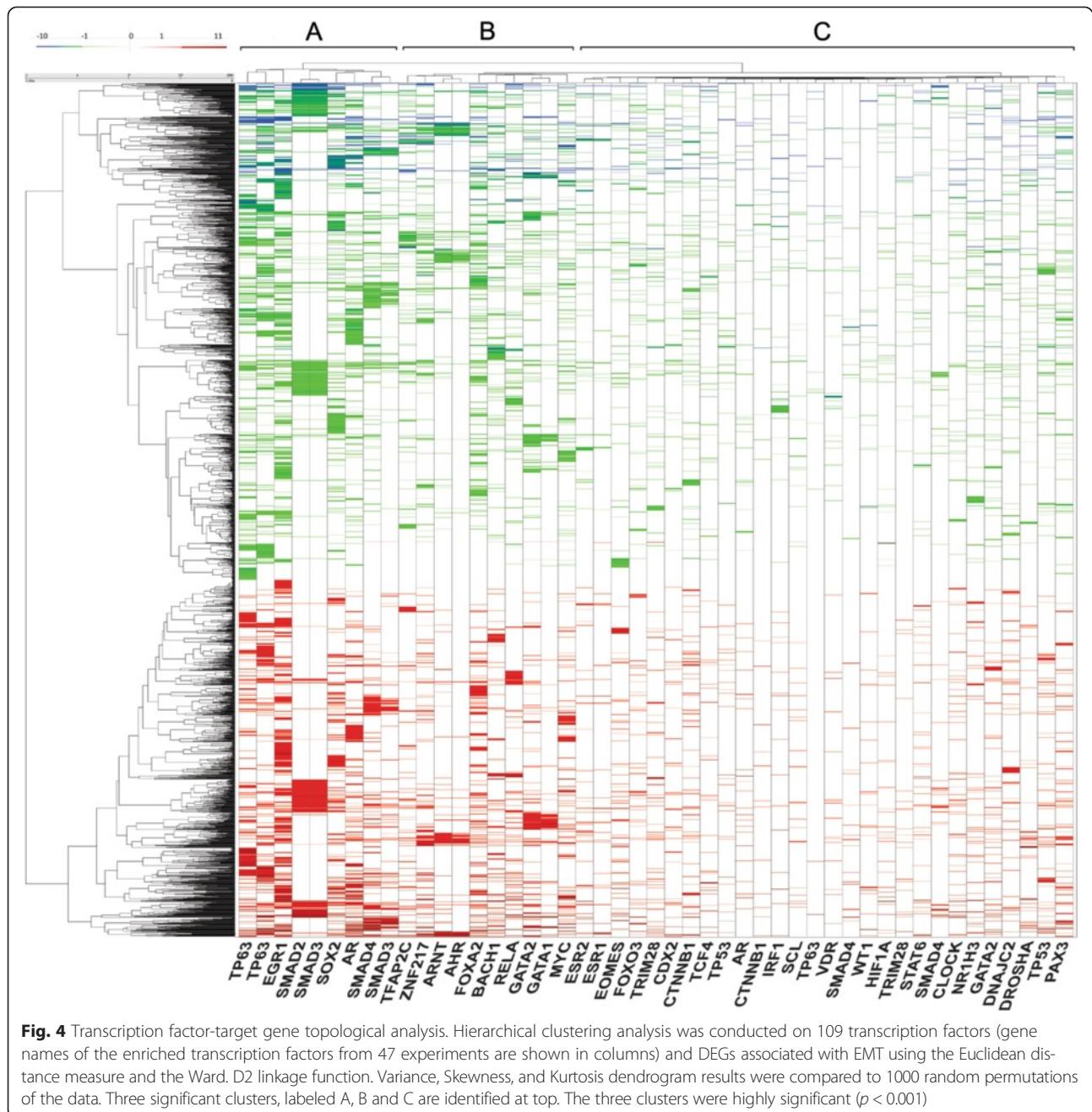
complexes referred to as the canonical and noncanonical pathways respectively [53, 54], producing sequential waves of NF- κ B dependent gene expression [55, 56]. Our mechanistic studies have found that the noncanonical pathway is coupled to the canonical pathway activation at a fixed time interval. This coupling is due to a feed-forward pathway mediated by inducible expression of the rate-limiting signaling adapter, TRAF1, that triggers activation of the noncanonical pathway [57]. Recently we observed that the EMT state produces profound changes in canonical-noncanonical pathway coupling through transcriptional reprogramming of TRAF1 and NF κ B2 promoters [5, 58]. As a result, the “coupling constant” e.g., the lag time between activation of the canonical and the noncanonical NF- κ B program, was markedly reduced by EMT.

We therefore hypothesized that TGF β -induced EMT produced profound differences in the response to activating the TNF signaling pathway. We first compared the TGF β -regulated gene network with that induced by TNF stimulation. Of the 3,487 DEGs detected in EMT, 547 were also regulated by TNF, indicating that the two signaling pathways are highly overlapping (Fig. 6a). The expression patterns of the TNF-dependent EMT regulators were examined. To reduce noise, the data set was filtered to only the more robustly

Table 7 Significantly depleted human TFs for the downregulated genes in the EMT process. The depleted transcription factors (TFs) were inferred by comparing ChIP-Seq or ChIP-chip identified target genes (TGs) of each TF with the downregulated genes in the EMT process. The depleted TFs were sorted according to the depleted fold ratios

Human TFs	PMID	# total TGs	# DETGs	Fold ratio	corrected <i>P</i> -value	Cell Lines
GABP	17652178	655	16	3.5	2.8E-09	JURKAT
KDM6A	18722178	410	14	2.5	2.2E-03	U937_AND_SAOS2
ETS1	20019798	1446	59	2.1	3.3E-10	JURKAT
EST1	17652178	636	28	2	3.3E-03	JURKAT
GABP	19822575	2430	132	1.6	4.4E-08	HepG2

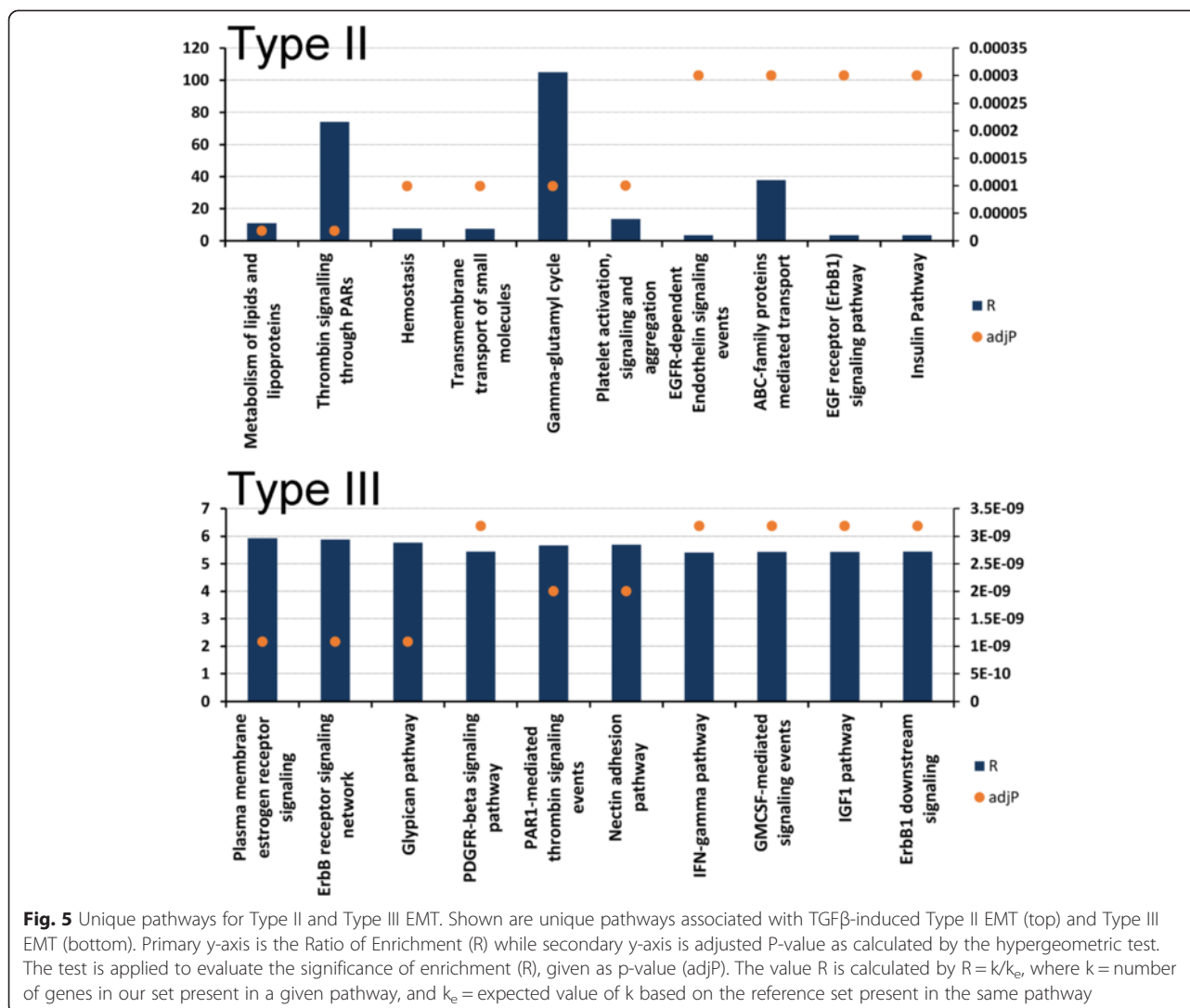
PMID pubmed article number; # total TGs number of ChIP-Seq or ChIP-chip identified target genes (TGs) of each TF; #DETGs number of differentially expressed genes



expressed 249 genes and subjected to hierarchical clustering, resulting in 6 major clusters (A-F, Fig. 6b). Cluster A represented early genes activated in hSAECs and silenced in EMT; Cluster B represented genes downregulated by TNF in hSAECs and not expressed in EMT; Cluster C represented late genes activated in hSAECs and whose expression was potentiated in EMT; Cluster D represented late genes activated in hSAECs and silenced in EMT; Cluster E represented early genes in hSAECs and potentiated by EMT; finally Cluster F represented genes not activated by TNF in hSAECs, but strongly upregulated by TNF in EMT in a late

expression pattern. TF enrichment analysis identified 13 shared transcription factors controlling these 6 clusters including androgen receptor (AR), GATA1/2, MYC, NANOG, NR0B1, NR1H3, RelA, SMAD2/3, STAT3, TP-53 and -63 (not shown). These are members from all three of the major TF clusters associated with the EMT program (Fig. 4).

The hierarchical clustering analysis suggested that the TNF pathway elicits distinct biological functions in epithelial cells versus that after Type II EMT. We therefore examined the function of genes activated by TNF in hSAECs but not activated after EMT, and those activated



after EMT but not in hSAECs. From this analysis, interferon signaling and immune system activation were prominent in hSAECs (Fig. 6c, top). By contrast, integrin signaling was upregulated by TNF in the EMT background (Fig. 6c, bottom). Integrin α V signaling is particularly important because other studies have demonstrated that integrin signaling drives the maintenance of the EMT state [59], providing an additional cross-talk mechanism for how cytokine signaling reinforces EMT.

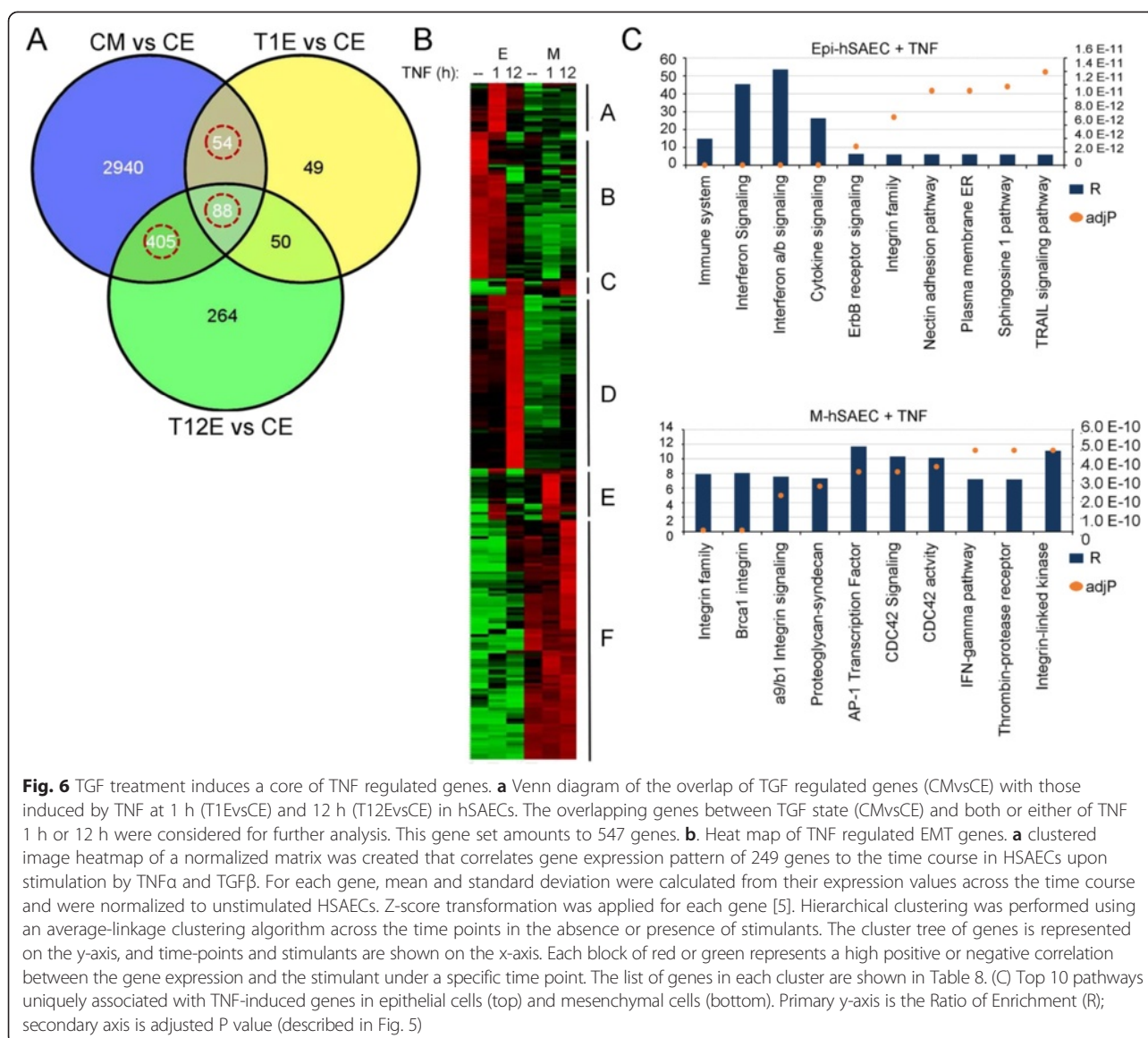
Analysis of TNF-dependent genes common to both epithelial and mesenchymal states, revealed that *NFKBIE/I κ B α* is found in Cluster E. *NFKBIE/I κ B α* is a prototypical early-response gene that we showed earlier was potentiated by EMT. In contrast, *TNFAIP1/Naf1*, a prototypic late gene under noncanonical NF- κ B pathway control is found in Cluster F consistent with our earlier observations [5, 55–57]. Together we conclude that TGF β -induced EMT produces complex modifications of

the TNF response program through both potentiation and inhibition of distinct gene subnetworks.

RelA is essential for TGF β -induced Type II EMT

Together, the findings that TGF β induces a gene expression profile similar to that of TNF (Fig. 2), a significant fraction of TGF β -regulated genes are NF- κ B target genes (Table 4), and that TGF β induces a core of TNF-induced NF- κ B-regulated genes (Fig. 6a) led us to investigate the relationship of NF- κ B activation with EMT.

Earlier studies have shown that TGF β is coupled to NF- κ B activation in carcinoma cells by oncogenic transformation induced expression of the TGF β -associated kinase, TAK1 [60]. Whether TGF β is coupled to NF- κ B activation in primary cells without TAK1 amplification is not known. We therefore measured the relative abundance of nuclear RelA in response to TGF β -stimulation. We observed a 2.5-fold induction of nuclear RelA 1 h after TGF β stimulation (Fig. 7a, top panel). We noted



that this induction was weaker than the 5-fold induction induced by the prototypical TNF α ligand; together there was no additive effect of the two ligands (Fig. 7a, lower panel).

Two strategies were used to determine the functional requirement of RelA in TGF β -induced EMT. First, the induction of EMT was examined in RelA $^{-/-}$ mouse embryonic fibroblasts (MEFs). Here, the expression of TGF β -induced *SNAIL1*, *IL-6* and *ZEB1* were completely blocked in RelA $^{-/-}$ MEFs compared to RelA WT MEFs (Fig. 7b). Second, we measured the effect of IKK inhibition on a time course experiment of TGF β stimulation in epithelial cells. For this purpose, we used a highly selective, allosteric inhibitor of IKK (BMS-345541); this compound does not affect the JNK, MAPK or Jak-STAT pathways [22]. We observed that TGF β stimulation

produced a complex pattern of *SNAIL1* expression, with a rapid peak of 30-fold induction after 2 days, followed by a decline, with a second peak of 40-fold induction after 5 days of treatment (Fig. 7c). By contrast, in epithelial cells treated with BMS-345541 at concentrations of 1, 3 and 10 μ M respectively, *SNAIL1* mRNA expression was significantly reduced at all time points measured. The patterns of *ZEB1* and *Twist1* showed a monotonic increase in expression peaking \sim 6 days after treatment at 40-fold and 3-fold inductions relative to untreated hSAECs, respectively. Similar to the effect on *SNAIL1*, BMS-345541 significantly reduced *ZEB1* expression and completely blocked *Twist1* expression (Fig. 7c). Similar patterns of inducible expression and BMS inhibition were observed for the mesenchymal cytoskeletal genes, *VIM*, *FNI* and *IL-6* (Fig. 7c). Together these data

Table 8 Gene list in each of the six clusters obtained from hierarchical clustering analysis of 249 genes upon stimulation by TNF α in epithelial and mesenchymal state. Each cluster corresponds to that in Fig. 6b

	Cluster A	Cluster B	Cluster C	Cluster D	Cluster E	Cluster F
1	RNF19B	DUSP6	SNCAIP	PIM3	PLAUR	SLC2A6
2	IRF1	EGR1	TNFAIP2	CLDN1	MAFF	SYNPO
3	GBP1	DUSP5	CSF1	SERPINB1	NKX3-1	TLR2
4	TNFSF15	SLITRK6	DRAM1	STARD5	IL6	COL12A1
5	BTG2	MAT2A	IKBKE	KRT6B	IL8	FEZ1
6	SPRR2D	NAB2	SLC39A8	GBP2	PTX3	PID1
7	HCAR3	CYP1A1		DHRS3	TNF	MGLL
8	HCAR2	FAM110C		SPRR1A	CXCL3	SERPINE2
9	IL1RN	KIAA1551		ITGB8	CXCL2	LBH
10	TNFAIP8	FBXW7		ANKRD33B	NFKBIZ	TNC
11	FOSB	SGK223		PI3	CSF2	MMP9
12	NR4A1	EMP1		CDC42EP4	IL1A	IL7R
13	NCOA7	SPRY2		LYPD3	IL1B	LAMB3
14	SERPINB2	TIPARP		LPAR6	NFKBIA	PIK3IP1
15	ARL5B	SRF		PRSS8	NFKBIE	NAMPT
16	ITPKC	TGIF1		DDX58	NUAK2	ITGA5
17	GATA6	FOSL1		C1orf226	LIF	DNER
18	SDC4	GADD45A		ISG15	TNFAIP3	GPR176
19		RAP2B		CXCL16	TNFAIP6	TGM2
20		OVOL1		SLPI		JPH2
21		PIM1		OAS1		COL27A1
22		HBEGF		TGM1		STC2
23		EREG		MUC1		FSTL3
24		TRIB1		PSTPIP2		SOX4
25		EGR3		SCNN1G		DUSP1
26		PHLDA2		SRD5A1		ST3GAL1
27		MAP3K14		S100A8		FADS3
28		LRIG1		MMP28		BHLHE40
29		KRT15		PRODH		RHOB
30		HAS3		IFIH1		SPRY4
31		IFFO2		ASS1		AKAP12
32		AKR1B10		GCLC		ANGPTL4
33		CKB		LRRC8D		DLC1
34		CALML3		OAS3		THBS1
35		SLC2A1		UBE2L6		SEMA7A
36		AQP3		IFI6		NNMT
37		IL6R		TAP1		ICAM1
38		SESN3		PSMB10		XBP1
39		ZNF462		HLA-F		CXCR7
40		DST		CDKN1C		MSC
41		HR		ALDH2		CXCL1
42		C1orf116		SQRDL		CSF3
43		WEE1		STAP2		ZC3H12A

Table 8 Gene list in each of the six clusters obtained from hierarchical clustering analysis of 249 genes upon stimulation by TNF α in epithelial and mesenchymal state. Each cluster corresponds to that in Fig. 6b (*Continued*)

44	GJB2	PRKAG2	INHBA
45	TSC22D1	SGPP2	SERPINA1
46	CYP1B1	C1R	EDN1
47	BCL10	TRANK1	G0S2
48	ALDH1A3	APOL6	LOC284454
49	PHLDA1	PARP14	BMP2
50	ERRFI1	CXCL10	DKK3
51	TCF4	BIRC3	DKK1
52	STK40	TYMP	SMAD7
53	AP5B1	CD74	PTPRE
54	FAM214A	SERPINA3	F2R
55		OAS2	TNFRSF12A
56		WFDC2	TNFRSF10D
57		SFTPD	ZBED2
58		RHCG	NEDD9
59		MX1	EMP3
60		PSMB9	GFPT2
61		BBOX1	MRAS
62		IFI27	AFAP1L1
63		SLC6A9	IL1R1
64		IFIT3	TLR4
65			MLPH
66			CCBE1
67			IGFL1
68			HRH1
69			AOX1
70			MARCKSL1
71			CHST2
72			SERPINE1
73			KCNU12
74			MCAM
75			PRSS23

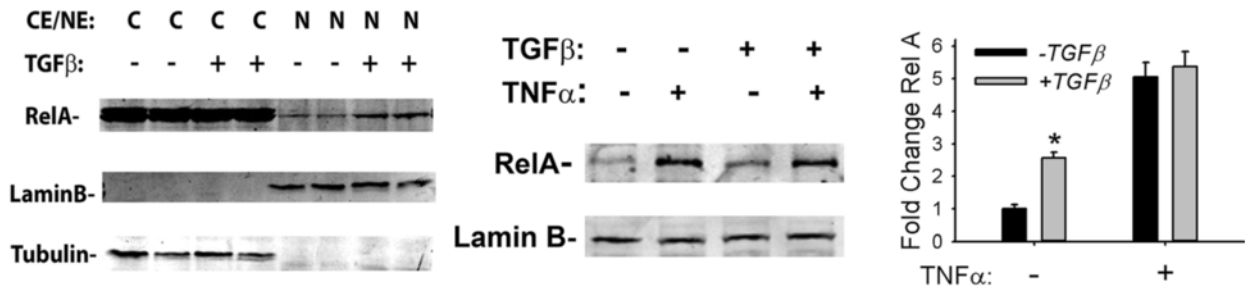
indicate that IKK-NF- κ B/RelA activation is a necessary component of the TGF β -induced EMT program.

Discussion

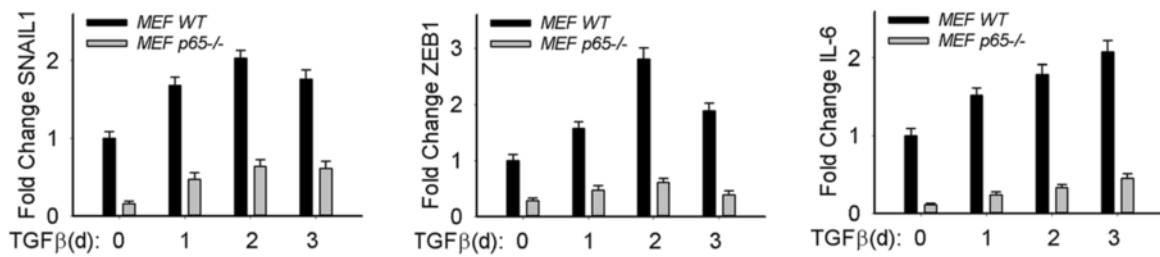
Type II EMT is a dynamic process that mediates airway response to injury, stimulating basal epithelial cells to promote re-epithelialization and extracellular matrix remodeling. EMT transcriptional reprogramming is a coordinated intracellular response triggered by epithelial growth factors (TGF, EGF) signaling via tyrosine kinase-coupled receptor growth factors to converge on a core set of transcription factors including SNAIL1, ZEB1, Twist1, and others. EMT is further modulated by

intercellular signaling cross-talk through innate inflammatory cytokines, extracellular matrix and ROS stress. In this study, we employed a well-established model of TGF β -induced Type II EMT in primary telomerase-immortalized hSAECs to understand the gene expression programs responsible for EMT in normal cells and how they are modified by the innate response. Surprisingly we observe that TGF β influences a 3,487 member gene network, 547 of which are also controlled by TNF. By comparison to genome-wide CHIP-Seq data sets, we observe that this core pathway is enriched in genes under NF- κ B/RelA control. We infer distinct biological functions compared to Type III EMT, demonstrate re-routing of the TNF

A



B



C

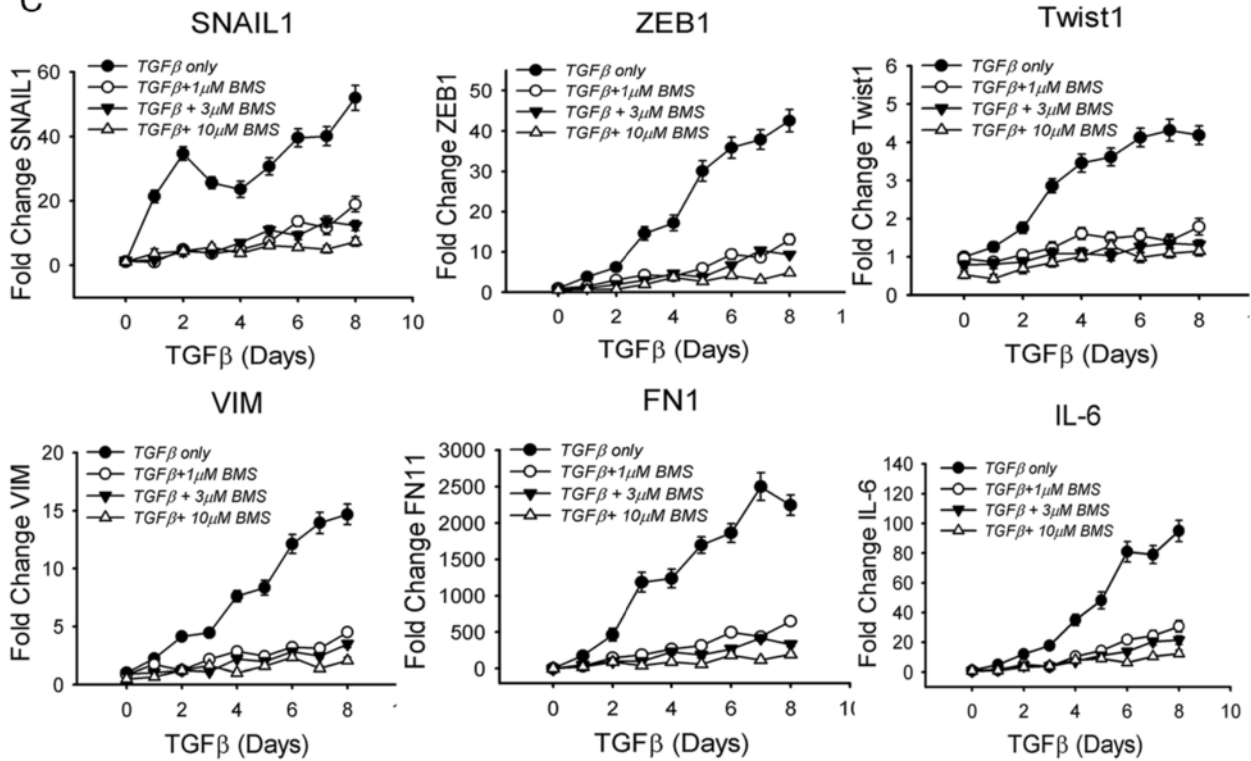


Fig. 7 (See legend on next page.)

(See figure on previous page.)

Fig. 7 Requirement of NF- κ B signaling for TGF β induced Type II EMT. **a** Cytoplasmic and nuclear extracts from hSAECs with or without EMT were isolated. 100 μ g of cytoplasmic or nuclear extracts were processed for Western blot using anti-RelA Ab (upper panel). Lamin B and tubulin were detected as loading control Abs respectively. Low left panel: epithelial cells or mesenchymal cells were stimulated with 25 μ g/ml TNF for 1 h and processed for RelA Western blot. Lower right panel: Quantification of nuclear RelA from Odyssey infrared imager. * $P < 0.05$ compared with control epithelial cells. **b** Wild-type (WT) or RelA^{-/-} mouse embryonic fibroblasts (MEF) were incubated with TGF β for 0, 1, 2, 3 days respectively. Total RNA was extracted and expression of EMT core transcription factors (*SNAIL1* and *ZEB1*) and NF- κ B-dependent genes (*IL-6*) were measured by Q-RT-PCR. The data shown is from n = 3 independent experiments. **c** Epithelial cells were stimulated with TGF β for various times in the absence or presence of a small molecule I κ B kinase inhibitor (BMS-345541, 1, 3, and 10 μ M respectively). The expression of *SNAIL1*, *Twist1*, *VIM*, *ZEB1*, *FN1* and *IL-6* were measured by Q-RT-PCR. The data shown are from n = 3 experiments

pathway and experimentally demonstrate the functional role of NF- κ B activation in Type II EMT. Finally, we note that the response of the TNF network is markedly affected by the EMT state, with distinct arms of the response being silenced, while others are enhanced. These studies are the first to our knowledge that demonstrate the unique TGF β signatures in inducing Type II EMT and the requirement of NF- κ B in this process.

TGF β -induced Type II EMT

TGF β activates cells by TGF β receptor type II (TGF β RII), a transmembrane serine- threonine kinase that recruits and phosphorylates TGF β RI on the cell membrane [7]. Activated TGF β RI triggers two inter-related pathways, known as the canonical (SMAD-dependent) and non-canonical (ERK/MAPK-dependent) pathways, that converge on a core set of transcription factors responsible for coordinated suppression of epithelial genes and induction of mesenchymal genes. In the canonical pathway, the activated TGF β RI/II complex phosphorylates cytoplasmic Smads 2/3 that bind to Smad4. This trimeric Smad complex binds to regulatory sequences in the E cadherin promoter, leading its repression and subsequent loss of adherens junctions [12]. In our study, we have found that *SNAIL1* expression is the most rapidly induced member of the core EMT regulators (Figs. 1b and 7c). Based on its rapid appearance, we suggest that *SNAIL1* expression is one of the early triggers of Type II EMT. The absence of *SNAIL1* in the public ChIP-Seq datasets prevented its enrichment and target gene analysis.

Other work has shown that the activated Smad 2/3/4 trimer also binds to Smad-binding elements in the regulatory regions of *JunB* and *c-Jun* genes, inducing their interaction with other coactivators including the EP300 coactivator [9]. This result indicates that the activities of the EMT core TFs are modulated by the EMT state, producing a bi-modal behavior, both activating and inhibiting gene expression. To this end, we note that over half of the TFs identified by enrichment analysis associated with upregulated DEGs in EMT are also associated with the downregulated genes in EMT. We note that these TFs include Smad2/3, RelA, SOX2 and others. Of relevance here, our earlier inference of RelA modulator-

target gene triplets showed 6 patterns of modulator interaction with RelA involved in activation and inhibition of specific target genes within biologically relevant pathways [35]. In extension of these studies, we predicted that *FN1* and *FOXPI* function as modulators of RelA action in the setting of EMT [61]. These latter predictions suggest that the target genes and behavior of RelA are influenced by the EMT. It will be of interest to develop systems-wide maps of the effects of TGF β on expression of modulators for the NF κ B/RelA and other EMT-regulated TFs.

Activated TGF β RI complex triggers a Smad-independent (“noncanonical”) pathway through the PI3K/Akt, Ras small GTPases, Wnt/ β -catenin, ERK, p38, and JNK kinases [7]. Although the actions of Smads are dominant on EMT, signaling through the noncanonical pathway is required for the full expression of EMT. Recent findings in non-small cell lung cancer cell lines indicate that phospho-Erk1/2 mediates a decrease in *ECad* and an increase in *FN1* expression [62]. Also, inhibition of JNK, p38 and Akt activities without affecting Smad phosphorylation blocks TGF β 1 induced α -*SMA*, *SNAIL1* and *col1A* in primary alveolar epithelial cells [63]. Together, these data indicate that Smad signaling is necessary but not sufficient for EMT. In the TF-TG topology map of EMT (Fig. 4), we note that the core TFs are clustered into three major clusters. We suggest that the TGs associated with Cluster A, containing Smad 2/3/4, TP63 and SOX2, are genes downstream of the canonical TGF β signaling pathway.

Transcription factors controlling Type II EMT

We computed the over-representation of 148 TF targets from publically available ChIP-chip, ChIP-Seq, ChIP-PET or DamID data [38] using a hypergeometric probability distribution to tentatively identify regulators of the DEGs in EMT. 18 of the 30 enriched transcription factors in the EMT upregulated genes overlapped with the 40 transcription factors regulating the EMT downregulated genes (p -value = 4.3e-13), indicating that these 18 had bimodal activities in a target gene dependent manner. Many of have been associated with EMT. In addition to Smad and TP63 proteins discussed above, NF-E2-related factor-2 (NFEL212/Nrf2) is negatively regulated by E-cadherin expression, whose downregulation

may promote Nrf2 nuclear translocation, resulting in the enhanced resistance of transformed cells to ROS stress [64]. The transcriptional repressor BACH1 has been shown to be upregulated in metastable Type III EMT [65]. We note that *RelA* and *BACH1*, *GATA1*, *GATA2* and *TFAP2C* are enriched in the upregulated genes, whereas only *TFAP2C* is enriched in the downregulated gene set. This observation suggests to us that the activation mode of RelA is modulated by association of distinct coactivators, discussed below. Further dissection of this transcription network will be of interest. For example, the implication of the histone acetyltransferase, CLOCK, in the EMT program and depletion of KDM6A have not been reported to our knowledge and merits further investigation.

TNF and TGF β cross-talk

TGF β -induced EMT is modulated by a variety of signals including cytokine stimulation, morphogen signaling, and ECM interactions. Of relevance here, studies of transformed alveolar basal epithelial cells show that TGF β -induced EMT is accelerated by the presence of members of the proinflammatory TNF/IL-1 superfamily of cytokines [16, 17]. TNF/IL-1 signal through highly conserved death domain containing receptors activating downstream Ras GTPase, p38 MAPK and JNK pathways, shared components of the noncanonical TGF β signaling pathway [7, 58]. By contrast to this modulatory role our observations in normal epithelial cells indicate that the TNF signature is a major component of TGF β -induced EMT (Fig. 5a). Pathway analysis of the DEGs in EMT indicates enrichment of TNFR signaling pathways (Fig. 3d) and we observe statistically significant enrichment of RelA target genes mapping to ChIP-Seq datasets (Table 4).

Requirement of NF- κ B in Type II EMT

NF- κ B is a master regulator of airway inflammation and cell fate determination. Our previous genome-wide analysis using whole genome-wide RNA-Seq and ChIP-Seq approaches have shown that this transcription factor regulates a ~4,000 member gene network mediating anti-apoptosis, inflammation, and adaptive immunity [34, 66]. In studies of cancer-associated Type III EMT, NF- κ B is required for IGF-induced EMT by directly inducing *SNAIL1* [67]. NF- κ B has also been shown to upregulate *ZEB1/2* and *Twist1* [20], explaining, in part, how the IL-1/TNF superfamily of cytokines mediates Type III EMT. Our studies indicate that *SNAIL1* is the earliest upregulated core transcription factor in Type II EMT and whose significant induction precedes that of *ZEB1* and *Twist1*, leading us to suggest that *SNAIL1* is an initial trigger of EMT.

Our study here shows that chronic TGF β stimulation induces nuclear translocation of NF- κ B/RelA. Other studies have shown that TGF β is a potent activator of NF- κ B/RelA in Ras-transformed cell lines. However, the relevance of these findings to normal epithelial cells is confounded by the observations that K-ras is a potent activator of NF- κ B [68], and that oncogenic transformation induces expression of the TGF β -associated kinase, TAK1, coupling NF- κ B to TGF β RI [60]. Our western blot quantification indicates that TGF β treatment in primary epithelial cells weakly induces NF- κ B/RelA translocation at levels significantly less than the prototypical TNF monokine (Fig. 7a).

NF- κ B/RelA translocation is controlled by two distinct pathways controlling sequential waves of NF- κ B dependent gene expression [55, 56]. The canonical pathway liberates cytoplasmic NF- κ B/RelA from I κ B α complexes in response to IKK α -IKK β [54], whereas the noncanonical cross-talk pathway liberates cytoplasmic NF- κ B/RelA from NF κ B2 complexes in response to IKK α -NIK. Our experiments using various doses of the allosteric IKK inhibitor indicate that expression of the core EMT genes is significantly inhibited at very low concentrations of BMS-345541. At this 1 μ M concentration, BMS is relatively selective for IKK, over that of other related MAPKs [22]. This behavior, together with our observations that RelA knockout cells are also defective in TGF β -induced *SNAIL1* and *ZEB* expression, support our conclusion that the IKK-NF κ B pathway is required for type II EMT. More work will be required to identify whether the NF- κ B pathway activated by TGF β is via the canonical, noncanonical or a combination of the two pathways.

EMT induced alterations in the TNF signaling program

Our earlier genome-wide microarray studies discovered that tonic cellular stimulation by TNF produces sequential waves of NF- κ B-dependent gene expression, with each wave controlling distinct biological functions [58, 59]. The most rapidly inducible, "early" genes were a group encoding cytokines, important in paracrine signaling, whereas late genes were those involved in antigen processing and cell surface presentation of MHC I complexes [56]. Mechanistically, the early genes are under canonical pathway control, and the late genes by noncanonical pathways [55, 56]. Recently we observed that the EMT state produces both profound accentuation of early gene expression and changes in canonical-noncanonical pathway coupling resulting in shortening the coupling interval between them [5, 58]. Although the mechanisms are still under investigation, part of this effect was through transcriptional reprogramming of NF κ B-dependent promoters, including that of *TRAF1* and *NF κ B2* [5, 58].

In this study we observe that the EMT state markedly changes the TNF program in unexpected ways. Although a cluster of early and late genes (Clusters E, C, respectively Fig. 6b) are potentiated by EMT, our data suggest that this is not a uniform behavior. Interestingly, large subgroups of early and late genes (Clusters A, D, Fig. 6b) are largely silenced in the EMT state. Conversely, a large cluster of genes are activated in a late gene expression pattern in mesenchymal cells that is not activated in normal epithelial cells (Cluster F, Fig. 6b). Pathway analysis shows that the TNF program in normal hSAECs is coupled to interferon signaling and innate immunity. By contrast, the TNF program is re-routed by Type II EMT to elicit the integrin signaling program, an essential pathway controlling maintenance of EMT [59].

How might the TNF signaling pathway be affected so dramatically by EMT? One explanation for this complex behavior is through contextual effects of NF- κ B associated modulators. In a recent study, we conducted a probabilistic inference of modulator activity on RelA activity using experimentally determined protein-protein interactions integrated with ChIP-Seq and a compendium of microarray data. The action identified 8349 modulator-RelA-target gene triplets, whose target genes were regulated in six distinct expression modes [35]. Using this approach, a preliminary inference of EMT-regulated NF- κ B/RelA modulators identified *FOXP1*, a tumor suppressor gene involved in cardiac development; *FNI* and *CCDC80*, interacting gene involved in extracellular matrix formation; *ACTN1*, a gene important in formation of actin polymers; and *DOCK10*, a Rho GTPase associated with cytoskeletal motility, as key modulators of the NF- κ B pathway in EMT [61]. These functions may suggest how the EMT state may modulate the activity of the ubiquitous NF- κ B/RelA in a cell-state dependent manner. More investigation will be required to determine the effect of EMT on the expression of RelA modulators. Moreover, these data suggest that the biological consequences of activating the TNF program are distinct in the normal vs EMT state.

Differences in Type II vs III EMT gene expression programs

TGF β -induced Type II EMT leads to disruption of mucosal barrier function by inducing the loss of apical polarity, reduced epithelial cadherin and disruption of epithelial adherens junctions [6], express α -SMA stress fibers and intermediate filament VIM, to produce ECM through secretion of Col1A and FN1, and to increase expression of MMPs to promote airway remodeling. This cellular biology is a distinct biological phenomenon from that of Type III induced in a primary transformed cell background, associated with acquisition of motility and resistance to chemotherapeutics.

Conclusions

Analysis of the gene program triggered by TGF β has led to a new understanding of the Type II EMT process. First, we have identified a core set of TNF-inducible genes enriched in experimentally determined NF- κ B/RelA binding sites. Second, we demonstrate that activation of NF- κ B/RelA is required for initiation of TGF β -induced core EMT TFs and mesenchymal genes. Third, we provide evidence that the gene program induced by EMT in normal epithelial cells confers distinct biological properties than that induced in mesenchymal cells. Finally, we demonstrate that the evolution of the TNF pathway is distinct in the EMT state, altering innate immunity and reinforcing EMT via integrin α V signaling. These data have implications for the effect of EMT on airway mucosal response in chronic respiratory disease.

Availability of supporting data

The RNA-Seq sequence data have been deposited in the NCBI GEO (Gene Expression Omnibus) database under the accession number GSE61220.

Additional files

Additional file 1: QA/QC of RNA Seq.

Additional file 2: Result of GLM for comparison of DEGs in E vs M state (no TGF β treatment vs plus TGF β).

Additional file 3: Result of GLM for comparison of DEGs in E state: 1 h TNF vs Control.

Additional file 4: Result of GLM for comparison of DEGs in E state 12 h TNF vs Control.

Additional file 5: Result of GLM for comparison of DEGs in E state 12 h TNF vs 1 h TNF.

Additional file 6: Result of GLM for comparison of DEGs after 1 h TNF E state vs 1 h TNF M state.

Additional file 7: Result of GLM for comparison of DEGs in 12 h TNF E state vs 12 h TNF M state.

Additional file 8: Result of GLM for comparison of DEGs in M state: 1 h TNF vs Control.

Additional file 9: Result of GLM for comparison of DEGs in M state 12 h TNF vs Control.

Additional file 10: Result of GLM for comparison of DEGs in M state 12 h TNF vs 1 h TNF.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

BT- developed cell system, conducted biological assays, wrote manuscript. JY- conducted biological assays, wrote manuscript. ARB- experimental design, wrote manuscript. SGW-NGS sequencing, analysis, wrote manuscript. BAL- modeling of NGS data, bioinformatics, wrote manuscript. TGW- NGS sequencing, analysis, wrote manuscript. MK, AK, XL, SB, BD- bioinformatics, wrote manuscript. All authors read and approved the final manuscript.

Authors' information

Bing Tian, Xueling Li, Mridul Kalita, Steven G. Widen Joint first authors
Thomas G. Wood, Bruce A. Luxon, Allan R. Brasier Joint senior authors

Acknowledgements

Research support was provided by funds from the Sealy Center for Molecular Medicine, UL1TR000071 UTMB CTSA (ARB), NIEHS P30 ES006676 (ARB), Keck Center for Interdisciplinary Bioscience Training of the Gulf Coast Consortium (CPRIT Grant No. RP140113, PI - Rathindra Bose, co-PI - B. Montgomery Pettitt, to XL, AK and ARB) and NSF/DMS Collaborative Grant No. 1361411 (to ARB, M. Kimmel and H. Levine).

Author details

¹Department of Internal Medicine, University of Texas Medical Branch (UTMB), Galveston, TX, USA. ²Sealy Center for Molecular Medicine, UTMB, Galveston, TX, USA. ³Institute for Translational Sciences, UTMB, Galveston, TX, USA. ⁴Department of Biochemistry and Molecular Biology, UTMB, Galveston, TX, USA. ⁵Bioinformatics Program, UTMB, Galveston, TX, USA.

Received: 12 September 2014 Accepted: 17 June 2015

Published online: 18 July 2015

References

- Knight DA, Holgate ST. The airway epithelium: structural and functional properties in health and disease. *Respirology*. 2003;8(4):432–46.
- Bals R, Hiemstra PS. Innate immunity in the lung: how epithelial cells fight against respiratory pathogens. *Eur Respir J*. 2004;23(2):327–33.
- Hippensteil S, Opitz B, Schmeck B, Suttorp N. Lung epithelium as a sentinel and effector system in pneumonia - molecular mechanisms of pathogen recognition and signal transduction. *Respir Res*. 2002;7(1):97.
- Kalluri R, Weinberg RA. The basics of epithelial-mesenchymal transition. *J Clin Invest*. 2009;119(6):1420–8.
- Kalita M, Tian B, Gao B, Choudhary S, Wood TG, Carmical JR, Boldogh I, Mitra S, Minna JD, Brasier AR: Systems approaches to modeling chronic mucosal inflammation. *BioMed Research International* 2013, in press (Application of Systems Biology and Bioinformatics Methods in Biochemistry and Biomedicine)
- Lambrecht BN, Hammad H. The airway epithelium in asthma. *Nat Med*. 2012;18(5):684–92.
- Willis BC, Borok Z. TGF-beta-induced EMT: mechanisms and implications for fibrotic lung disease. *Am J Physiol Lung Cell Mol Physiol*. 2007;293(3):L525–34.
- Katsuno Y, Lamouille S, Derynck R. TGF-beta signaling and epithelial-mesenchymal transition in cancer progression. *Curr Opin Oncol*. 2013;25(1):76–84.
- Boxall C, Holgate ST, Davies DE. The contribution of transforming growth factor-beta and epidermal growth factor signalling to airway remodelling in chronic asthma. *Eur Respir J: Off J Euro Soc Clin Res Physiol*. 2006;27(1):208–29.
- Herranz N, Pasini D, Diaz VM, Franci C, Gutierrez A, Dave N, et al. Polycomb complex 2 is required for E-cadherin repression by the Snail1 transcription factor. *Mol Cell Biol*. 2008;28(15):4772–81.
- Battle E, Sancho E, Franci C, Dominguez D, Monfar M, Baulida J, et al. The transcription factor snail is a repressor of E-cadherin gene expression in epithelial tumour cells. *Nat Cell Biol*. 2000;2(2):84–9.
- Vincent T, Neve EP, Johnson JR, Kukalev A, Rojo F, Albanell J, et al. A SNAIL1-SMAD3/4 transcriptional repressor complex promotes TGF-beta mediated epithelial-mesenchymal transition. *Nat Cell Biol*. 2009;11(8):943–50.
- McDonald OG, Wu H, Timp W, Doi A, Feinberg AP. Genome-scale epigenetic reprogramming during epithelial-to-mesenchymal transition. *Nat Struct Mol Biol*. 2011;18(8):867–74.
- Wang J, Scully K, Zhu X, Cai L, Zhang J, Prefontaine GG, et al. Opposing LSD1 complexes function in developmental gene activation and repression programmes. *Nature*. 2007;446(7138):882–7.
- Lamouille S, Xu J, Derynck R. Molecular mechanisms of epithelial-mesenchymal transition. *Nat Rev Mol Cell Biol*. 2014;15(3):178–96.
- Camara J, Jarai G. Epithelial-mesenchymal transition in primary human bronchial epithelial cells is Smad-dependent and enhanced by fibronectin and TNF-alpha. *Fibrogenesis Tissue Repair*. 2010;3(1):2.
- Kasai H, Allen JT, Mason RM, Kamimura T, Zhang Z. TGF-beta1 induces human alveolar epithelial to mesenchymal cell transition (EMT). *Respir Res*. 2005;6:56.
- Hsu H, Shu HB, Pan MG, Goeddel DV. TRADD-TRAF2 and TRADD-FADD interactions define two distinct TNF receptor 1 signal transduction pathways. *Cell*. 1996;84:299–308.
- Liu ZG, Hsu H, Goeddel DV, Karin M. Dissection of TNF receptor 1 effector functions: JNK activation is not linked to apoptosis while NF-kB activation prevents cell death. *Cell*. 1996;87:565–76.
- Li C-W, Xia W, Huo L, Lim S-O, Wu Y, Hsu JL, et al. Epithelial-mesenchymal transition induced by TNF-alpha requires NF-kB-mediated transcriptional upregulation of Twist1. *Cancer Res*. 2012;72(5):1290–300.
- Ramirez RD, Sheridan S, Girard L, Sato M, Kim Y, Pollack J, et al. Immortalization of human bronchial epithelial cells in the absence of viral oncoproteins. *Cancer Res*. 2004;64(24):9027–34.
- Burke JR, Pattoli MA, Gregor KR, Brassil PJ, MacMaster JF, McIntyre KW, et al. BMS-345541 Is a Highly Selective Inhibitor of IκB Kinase That Binds at an Allosteric Site of the Enzyme and Blocks NF-kB-dependent Transcription in Mice. *J Biol Chem*. 2003;278(3):1450–6.
- Tian B, Zhao Y, Kalita M, Edeh CB, Paessler S, Casola A, et al. CDK9-dependent transcriptional elongation in the innate interferon-stimulated gene response to respiratory syncytial virus infection in airway epithelial cells. *J Virol*. 2013;87(12):7075–92.
- Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2^{-ΔΔC_T} Method. *Methods*. 2001;25(4):402–8.
- Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2010;26(1):139–40.
- FastQC: A quality control tool for high throughput sequence data. [<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>]
- Kim D, Perteau G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol*. 2013;14(4):R36.
- Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012;9(4):357–9.
- Benjamini Y, Hochberg Y. *J R Stat Soc Ser B Methodol*. 1995;57(1):298–300.
- Dunn OJ. Multiple comparisons among means. *J Am Stat Assoc*. 1961;56(293):52–64.
- Lachmann A, Xu H, Krishnan J, Berger SI, Mazloom AR, Ma'ayan A. ChEA: transcription factor regulation inferred from integrating genome-wide ChIP-X experiments. *Bioinformatics*. 2010;26(19):2438–44.
- Huang W, Loganathanaraj R, Schroeder B, Fargo D, Li L. PAVIS: a tool for peak annotation and visualization. *Bioinformatics*. 2013;29(23):3097–9.
- Huang RY, Guilford P, Thiery JP. Early events in cell adhesion and polarity during epithelial-mesenchymal transition. *J Cell Sci*. 2012;125(Pt 19):4417–22.
- Yang J, Mitra A, Dojer N, Fu S, Rowicka M, Brasier AR: A probabilistic approach to learn chromatin architecture and accurate inference of the NF-kappaB/RelA regulatory network using ChIP-Seq. *Nucleic acids research* 2013;41(15):7240–59.
- Li X, Zhao Y, Tian B, Jamaluddin M, Mitra A, Yang J, et al. Modulation of gene expression regulated by the transcription factor NF-kappaB/RelA. *J Biol Chem*. 2014;289(17):1927–44.
- Chen L, Munoz-Antonia T, Cress WD. Trim28 contributes to EMT via regulation of E-cadherin and N-cadherin in lung cancer cell lines. *PLoS One*. 2014;9(7), e101040.
- Li X, Xu Y, Chen Y, Chen S, Jia X, Sun T, et al. SOX2 promotes tumor metastasis by stimulating epithelial-to-mesenchymal transition via regulation of WNT/beta-catenin signal network. *Cancer Lett*. 2013;336(2):379–89.
- Wang X, Belguise K, Kersual N, Kirsch KH, Mineva ND, Galtier F, et al. Oestrogen signalling inhibits invasive phenotype by repressing RelB and its target BCL2. *Nat Cell Biol*. 2007;9(4):470–8.
- Mak P, Leav I, Pursell B, Bae D, Yang X, Taglienti CA, et al. ERbeta impedes prostate cancer EMT by destabilizing HIF-1alpha and inhibiting VEGF-mediated snail nuclear localization: implications for Gleason grading. *Cancer Cell*. 2010;17(4):319–32.
- Hwang-Verslues WW, Chang PH, Jeng YM, Kuo WH, Chiang PH, Chang YC, et al. Loss of corepressor PER2 under hypoxia up-regulates OCT1-mediated EMT gene expression and enhances tumor malignancy. *Proc Natl Acad Sci U S A*. 2013;110(30):12331–6.
- Chang CJ, Chao CH, Xia W, Yang JY, Xiong Y, Li CW, et al. p53 regulates epithelial-mesenchymal transition and stem cell properties through modulating miRNAs. *Nat Cell Biol*. 2011;13(3):317–23.

42. Wang Z, Jiang Y, Guan D, Li J, Yin H, Pan Y, et al. Critical roles of p53 in epithelial-mesenchymal transition and metastasis of hepatocellular carcinoma cells. *PLoS One*. 2013;8(9), e72846.
43. Olsen JR, Oyan AM, Rostad K, Hellem MR, Liu J, Li L, et al. p63 attenuates epithelial to mesenchymal potential in an experimental prostate cell model. *PLoS One*. 2013;8(5), e62547.
44. Tran MN, Choi W, Wszolek MF, Navai N, Lee IL, Nitti G, et al. The p63 protein isoform DeltaNp63alpha inhibits epithelial-mesenchymal transition in human bladder cancer cells: role of MIR-205. *J Biol Chem*. 2013;288(5):3275–88.
45. Liu X, Ru J, Zhang J, Zhu LH, Liu M, Li X, et al. miR-23a targets interferon regulatory factor 1 and modulates cellular proliferation and paclitaxel-induced apoptosis in gastric adenocarcinoma cells. *PLoS One*. 2013;8(6), e64707.
46. Cyr AR, Kulak MV, Park JM, Bogachek MV, Spanheimer PM, Woodfield GW, White-Baer LS, O'Malley YQ, Sugg SL, Olivier AK et al.: TFAP2C governs the luminal epithelial phenotype in mammary development and carcinogenesis. *Oncogene* 2014;34(4):436–44.
47. Rico-Leo EM, Alvarez-Barrientos A, Fernandez-Salguero PM. Dioxin receptor expression inhibits basal and transforming growth factor beta-induced epithelial-to-mesenchymal transition. *J Biol Chem*. 2013;288(11):7841–56.
48. Chakrabarti R, Hwang J, Andres Blanco M, Wei Y, Lukacisin M, Romano RA, et al. E1f5 inhibits the epithelial-mesenchymal transition in mammary gland development and breast cancer metastasis by transcriptionally repressing Snail2. *Nat Cell Biol*. 2012;14(11):1212–22.
49. Stewart CA, Wang Y, Bonilla-Claudio M, Martin JF, Gonzalez G, Taketo MM, et al. CTNNB1 in mesenchyme regulates epithelial cell differentiation during Mullerian duct and postnatal uterine development. *Mol Endocrinol*. 2013;27(9):1442–54.
50. Johnson L, Mercer K, Greenbaum D, Bronson RT, Crowley D, Tuveson DA, et al. Somatic activation of the K-ras oncogene causes early onset lung cancer in mice. *Nature*. 2001;410(6832):1111–6.
51. Foster KA, Oster CG, Mayer MM, Avery ML, Audus KL. Characterization of the A549 cell line as a type II pulmonary epithelial cell model for drug metabolism. *Exp Cell Res*. 1998;243(2):359–66.
52. Lieber M, Smith B, Szakal A, Nelson-Rees W, Todaro G. A continuous tumor-cell line from a human lung carcinoma with properties of type II alveolar epithelial cells. *Int J Cancer*. 1976;17(1):62–70.
53. Brasier AR: The NF- κ B Signaling Network: Insights from systems approaches. In: *Cellular Signaling And Innate Immune Responses To RNA Virus Infections*. Edited by Brasier AR, Lemon SM, Garcia-Sastre A: American Society for Microbiology Press, Herndon, VA. 2008: 119–135.
54. Brasier AR. The NF- κ B regulatory network. *Cardiovasc Toxicol*. 2006;6(2):111–30.
55. Nowak DE, Tian B, Jamaluddin M, Boldogh I, Vergara LA, Choudhary S, et al. RelA Ser276 phosphorylation is required for activation of a subset of NF- κ B-dependent genes by recruiting cyclin-dependent kinase 9/cyclin T1 complexes. *Mol Cell Biol*. 2008;28(11):3623–38.
56. Tian B, Nowak DE, Brasier AR. A TNF-induced gene expression program under oscillatory NF- κ B control. *BMC Genomics*. 2005;6:137–54.
57. Choudhary S, Kalita M, Fang L, Patel K, Tian B, Zhao Y, et al. Inducible TNF receptor associated factor 1 expression couples the canonical to the Non-canonical NF- κ B pathway in TNF stimulation. *J Biol Chem*. 2013;288(20):14612–23.
58. Ijaz T, Pazdrak K, Kalita M, Konig R, Choudhary S, Tian B, et al. Systems biology approaches to understanding epithelial mesenchymal transition (EMT) in mucosal remodeling and signaling in asthma. *World Allergy Organ J*. 2014;7(1):13.
59. Mamuya FA, Duncan MK. α V integrins and TGF-beta-induced EMT: a circle of regulation. *J Cell Mol Med*. 2012;16(3):445–55.
60. Freudlsperger C, Bian Y, Contag Wise S, Burnett J, Coupar J, Yang X, et al. TGF- β and NF- κ B signal pathway cross-talk is mediated through TAK1 and SMAD7 in a subset of head and neck cancers. *Oncogene*. 2013;32(12):1549–59.
61. Li X, Zhu M, Brasier AR, Kudlicki AS. Inferring genome-wide functional modulatory network: a case study on NF- κ B/RelA transcription factor. *J Comp Bio: J Comp Mol Cell Bio*. 2015;22(4):300–12.
62. Buonato JM, Lazzara MJ. ERK1/2 blockade prevents epithelial-mesenchymal transition in lung cancer cells and promotes their sensitivity to EGFR inhibition. *Cancer Res*. 2014;74(1):309–19.
63. Xi Y, Tan K, Brumwell AN, Chen SC, Kim YH, Kim TJ, et al. Inhibition of epithelial-to-mesenchymal transition and pulmonary fibrosis by methacycline. *Am J Respir Cell Mol Biol*. 2014;50(1):51–60.
64. Kim WD, Kim YW, Cho JJ, Lee CH, Kim SG. E-cadherin inhibits nuclear accumulation of Nrf2: implications for chemoresistance of cancer cells. *J Cell Sci*. 2012;125(Pt 5):1284–95.
65. Thomson S, Petti F, Sujka-Kwok I, Mercado P, Bean J, Monaghan M, et al. A systems view of epithelial-mesenchymal transition signaling states. *Clin Exp Metastasis*. 2011;28(2):137–55.
66. Tian B, Brasier AR. Identification of an NF- κ B dependent gene network. *Recent Prog Horm Res*. 2003;58:95–130.
67. Kim HJ, Litzenburger BC, Cui X, Delgado DA, Grabner BC, Lin X, et al. Constitutively active type I insulin-like growth factor receptor causes transformation and xenograft growth of immortalized mammary epithelial cells and is accompanied by an epithelial-to-mesenchymal transition mediated by NF- κ B and snail. *Mol Cell Biol*. 2007;27(8):3165–75.
68. Bassères DS, Ebbs A, Levantini E, Baldwin AS. Requirement of the NF- κ B subunit p65/RelA for K-Ras-induced lung tumorigenesis. *Cancer Res*. 2010;70(9):3537–46.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

