



OPEN

Integrated bioinformatics analysis identifies established and novel TGFβ1-regulated genes modulated by anti-fibrotic drugs

Ava C. Wilson^{1,2}, Joe Chiles², Shah Ashish³, Diptiman Chanda², Preeti L. Kumar², James A. Mobley⁴, Enid R. Neptune⁵, Victor J. Thannickal^{2,6} & Merry-Lynn N. McDonald^{1,2,7}✉

Fibrosis is a leading cause of morbidity and mortality worldwide. Although fibrosis may involve different organ systems, transforming growth factor-β (TGFβ) has been established as a master regulator of fibrosis across organs. Pirfenidone and Nintedanib are the only currently-approved drugs to treat fibrosis, specifically idiopathic pulmonary fibrosis, but their mechanisms of action remain poorly understood. To identify novel drug targets and uncover potential mechanisms by which these drugs attenuate fibrosis, we performed an integrative 'omics analysis of transcriptomic and proteomic responses to TGFβ1-stimulated lung fibroblasts. Significant findings were annotated as associated with pirfenidone and nintedanib treatment in silico via Coremine. Integrative 'omics identified a co-expressed transcriptomic and proteomic module significantly correlated with TGFβ1 treatment that was enriched (FDR-p = 0.04) with genes associated with pirfenidone and nintedanib treatment. While a subset of genes in this module have been implicated in fibrogenesis, several novel TGFβ1 signaling targets were identified. Specifically, four genes (*BASP1*, *HSD17B6*, *CDH11*, and *TNSI*) have been associated with pirfenidone, while five genes (*CLINT1*, *CADM1*, *MTDH*, *SYDE1*, and *MCTS1*) have been associated with nintedanib, and *MYDGF* has been implicated with treatment using both drugs. Using the Clue Drug Repurposing Hub, succinic acid was highlighted as a metabolite regulated by the protein encoded by *HSD17B6*. This study provides new insights into the anti-fibrotic actions of pirfenidone and nintedanib and identifies novel targets for future mechanistic studies.

Affecting all organs and contributing to numerous diseases, fibrosis is a leading cause of mortality and morbidity, accounting for up to 45% of all deaths globally¹. Fibrosis is caused by abnormal extracellular matrix (ECM) deposition by activated (myo)fibroblasts that results in scarring of organ-specific tissues^{2–4}. Transforming growth factor beta 1 (TGFβ1) plays a central role in fibrogenesis, primarily by inducing the transition of fibroblasts into myofibroblasts⁵. New candidates for treating fibrotic diseases are needed as only two drugs (pirfenidone and nintedanib) have received regulatory approval to treat fibrosis, specifically idiopathic pulmonary fibrosis (IPF)⁶. IPF is an age-related interstitial pulmonary disease characterized by worsening dyspnea and reduced lung function due to progressive, irreversible fibrosis^{7,8}. In addition to a poor quality of life, the prognosis for IPF patients is poor with a median survival time of 3 years after diagnosis⁹. Although both pirfenidone and nintedanib are widely used in the treatment of IPF the underlying mechanisms of their biological action(s) remains poorly understood^{10–12}. Further, pirfenidone and nintedanib have varying efficacy in treating IPF necessitating a deeper understanding of their mechanisms in delaying disease progression^{13,14}.

Currently, there is limited data regarding the molecular mechanisms of pirfenidone and nintedanib on the canonical TGFβ1 pathway. Reports of the potential anti-fibrotic actions of pirfenidone have been limited to

¹Department of Epidemiology, School of Public Health, University of Alabama at Birmingham, Birmingham, AL, USA. ²Division of Pulmonary, Allergy and Critical Care Medicine, Department of Medicine, University of Alabama at Birmingham, Birmingham, AL, USA. ³Department of Orthopedic Surgery, University of Alabama at Birmingham, Birmingham, AL, USA. ⁴Division of Molecular and Translational Biomedicine, Department of Anesthesiology and Perioperative Medicine, University of Alabama at Birmingham, Birmingham, AL, USA. ⁵Department of Medicine, Johns Hopkins University, Baltimore, MD, USA. ⁶John W. Deming Department of Medicine, Tulane University School of Medicine, New Orleans, LA, USA. ⁷Department of Genetics, University of Alabama at Birmingham, Birmingham, AL, USA. ✉email: mldonnely@uabmc.edu

candidate gene analyses¹⁵. The efficacy of pirfenidone in treating fibrosis has been attributed to pleiotropic modes of action including anti-fibrotic, anti-inflammatory, and antioxidant effects¹⁶. Ballester et al. found pirfenidone's inhibitory effect on TGFβ1-induced fibrosis is mediated through various mechanisms including the inhibition of transmembrane mucin 1 c-terminal cytoplasmic tail (MUC1-CT) phosphorylation, β-catenin activation, nuclear complex formation of the phospho-SMAD3/MUC1-CT/active-β-catenin complex, and SMAD-binding element (SBE) activation¹⁵. Nintedanib is a tyrosine kinase inhibitor that inhibits multiple tyrosine kinase signaling pathways including VEGF, PDGF, and FGF¹⁴. Both drugs have pleiotropic effects targeting different fibrotic mechanisms, however, multi-omics studies investigating the antifibrotic mechanism of pirfenidone and nintedanib in relation to TGFβ1 signaling responses are lacking. Although both drugs slow disease progression, neither drug improves or stabilizes lung function, nor improves quality of life¹⁷. In addition, both drugs have tolerability issues, facilitating the need for the identification of novel IPF drug targets¹⁷.

To begin to address these gaps in our understanding of fibrotic disease development and treatment, we hypothesized that identifying TGFβ1 induced transcriptomic and proteomic changes that are co-expressed with genes associated with response to pirfenidone and nintedanib may identify novel mechanisms of drug action, and potentially uncover new targets for future anti-fibrotic therapies. To test this hypothesis, transcriptomic and proteomic profiles with and without TGFβ1 treatment of human fetal lung mesenchymal cells (IMR-90) were analyzed. TGFβ1-induced changes integrated across the transcriptome and proteome were identified at the single transcript and protein and network levels using Weighted Gene Co-Expression Network Analysis (WGCNA) software¹⁸. Further, we provide a robust annotation of integrated transcriptomic and proteomic signals significantly correlated with TGFβ1 treatment, including mining of the literature for pirfenidone and nintedanib associated genes, mining of known functional pathways to identify TGFβ1 signaling targets, screening of drug repurposing metadata, and pathway analyses.

Results

Differentially expressed genes and proteins induced by TGFβ1. At the single gene level, TGFβ1 induced differential expression of 780 genes (Table S1). Among these, 416 genes were upregulated, and 364 genes were downregulated (Table S1). Collectively, the 780 TGFβ1 differentially expressed genes were enriched with genes involved in collagen deposition in the extracellular matrix (ECM) (GO Collagen Containing Extracellular Matrix, FDR p-value = 3.84×10^{-34}), in addition to genes involved with apoptosis (GO Apoptotic Process, FDR p-value = 1.89×10^{-43}), extracellular signaling (GO Enzyme Linked Receptor Protein Signaling Pathway, FDR p-value = 1.22×10^{-38}), and genes involved in the epithelial mesenchymal transition (EMT) in wound healing and fibrosis (Hallmark Epithelial Mesenchymal Transition, FDR p-value = 3.31×10^{-79}) (Table S2).

At the single protein level, TGFβ1 significantly altered the levels of eight proteins (Table S3). Of these proteins, TGFβ1 stimulated an increase for six proteins (PDZ and LIM domain protein 5, Calponin 1, Collagen alpha-1 V chain, Tensin 1, Calponin 3, and LIM domain and actin binding protein 1) and a decrease for two proteins (Calpain 2 catalytic subunit and Collagen alpha-1 IV chain 1) (Table S3). Collectively this set of TGFβ1 altered proteins is enriched with those encoding ECM proteoglycans (REACTOME ECM Proteoglycans, FDR p-value = 0.032), collagen biosynthesis and modifying enzymes (REACTOME Collagen Biosynthesis and Modifying Enzymes, FDR p-value = 0.0281), and genes up-regulated by TGFβ1 (McBryan Pubertal TGF-β1 Targets Up, FDR p-value = 0.00447) (Table S4).

TGFβ1 induced co-expression of transcriptomic modules. At the network level, weighted gene co-expression analyses identified 11 modules of co-expressed genes. TGFβ1 induced significant upregulation of the 1564 blue module transcripts ($r_{\text{blue}} = 1.0$, $p_{\text{blue}} = 1 \times 10^{-5}$, Fig. 1A, Table S5). In the blue transcriptomics module, 677 were novel genes not previously implicated in TGFβ1 signaling (Table S5). The mean and standard deviation of the kME for the blue module was 0.87 ± 0.14 with 5.3% of the transcripts meeting the hub criteria of $kME > 0.99$ (Table S6). Visualization of the connectivity of the blue module hub genes is depicted in Fig. 1B.

Network analyses indicates two proteomic modules correlated with TGFβ1 treatment. At the network level, the weighted gene co-expression protein network consisted of 7 modules. Of which, the yellow and turquoise modules were significantly correlated with TGFβ1 treatment at levels withstanding Bonferroni correction (Fig. 2A). Collectively, the 78 proteins in the yellow module ($r_{\text{yellow}} = -0.94$, $p_{\text{yellow}} = 0.006$) were significantly downregulated with TGFβ1 treatment (Fig. 2A). In the turquoise module, 115 proteins ($r_{\text{turquoise}} = 0.99$, $p_{\text{turquoise}} = 9 \times 10^{-5}$) module were collectively upregulated with TGFβ1 treatment (Fig. 2A). The mean and standard deviation of the kME for the yellow and turquoise protein modules were 0.78 ± 0.13 and 0.73 ± 0.18 , respectively (Table S6). Additionally, 23.1% and 21.7% of the proteins in the yellow and turquoise modules, respectively, were hub proteins (Table S6). Visualization of the connectivity of hub proteins, $kME > 0.9$, in the yellow (N = 18) and turquoise (N = 25) modules is depicted in Fig. 2B,C. Of the 78 proteins in the yellow module, 26 were identified as novel proteins involved in TGFβ1 signaling, with 2 proteins (*LGALS1* and *COL6A2*) known to be secreted proteins (Table S5). Of the 115 proteins in the turquoise module, 43 were identified as novel proteins involved in TGFβ1 signaling, with 5 proteins (*COPA*, *TGFBI*, *COL5A2*, *PXDND*, and *COL5A1*) known to be secreted (Table S5).

Integrative 'omics analyses identifies co-expressed transcript and peptides modules that correlate with TGFβ1 treatment. The blue transcriptomics module was correlated with the turquoise and yellow proteomic modules (Fig. 1C). The expression of genes in the blue module was positively correlated with the proteins in the turquoise module (Fig. 1C: $r = 0.98$, $p = 3 \times 10^{-4}$). Whereas the expression of genes in the blue module genes was negatively correlated with protein levels in the yellow module but at a nominal ($p < 0.05$) level

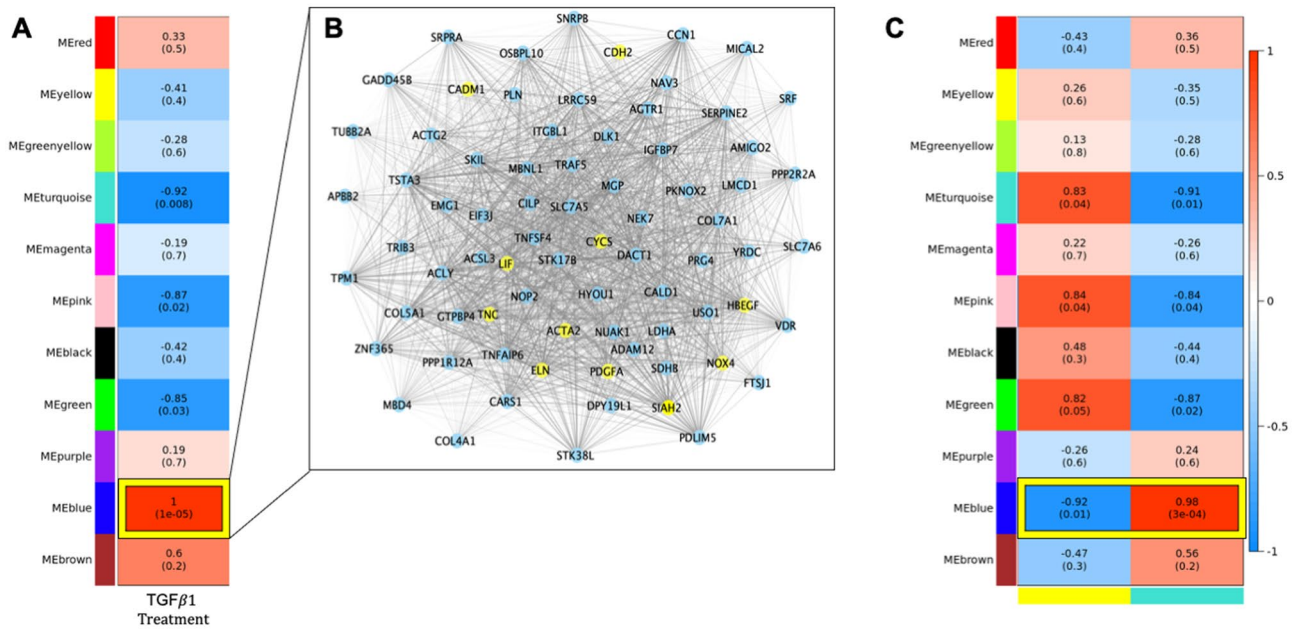


Figure 1. Results of integrative omics analysis of proteomic and transcriptomic data generated from IMR-90 cells with and without TGFβ1 treatment. **(A)** Transcriptomic module association with TGFβ1 treatment: Values in each cell represent correlation, in parentheses, with p-values between each module of co-expressed transcripts and TGFβ1 treatment. Heatmap shading corresponds to strength of association where darker red cells have higher upregulation and darker blue cells have higher downregulation based on correlation. Cells outlined in yellow withstand Bonferroni correction for multiple testing based on the number of modules generated. **(B)** Network visualization of hub genes in the blue transcriptomic module. Genes with a kME larger than 0.99 were selected for visualization in the blue module. The thickness of the edge corresponds to increasing topological overlap (TOM), a measure of the strength of correlation between transcript levels, which is the Pearson's correlation obtained from the adjacency matrix. Nodes labeled in yellow correspond to single genes in the blue module that are annotated as associated with pirfenidone and/or nintedanib treatment. **(C)** Results from integration of transcriptomic and proteomic data. Values in each cell represent correlation, p-values in parentheses, between each module of co-expressed transcripts with TGFβ1 treatment and modules of co-expressed proteins. The y-axis corresponds to transcriptomic modules generated using WGCNA. The x-axis corresponds to the yellow and turquoise proteomic modules. Individually, the yellow and turquoise proteomic modules were significantly correlated with TGFβ1 treatment (depicted in Fig. 2).

(Fig. 1C: $r = -0.92$, $p = 0.01$). Several of the TGFβ1 significantly differentially expressed genes (Table S5) had encoded proteins which were hub members of the yellow (*COL4A1*) and turquoise (*CNN1*, *TNS1*, *COL5A1*, and *PDLIM5*) proteomics module (Fig. 2B,C). The five genes and proteins were enriched for pro-fibrotic pathways including collagen formation (REACTOME collagen formation, FDR p-value = 0.021), ECM proteoglycans (REACTOME ECM proteoglycans, FDR p-value = 0.017), and PDGF signaling (REACTOME PDGF Signaling, FDR p-value = 0.016), among others.

Novel fibrotic biomarkers common to pirfenidone, nintedanib, and TGFβ1 signaling. Of the genes annotated as response to pirfenidone and/or nintedanib treatment, 264 genes were specific to pirfenidone and 150 genes were specific to nintedanib. A total of 149 genes were identified as being associated with both pirfenidone and nintedanib treatment (Table S8). The blue module was significantly enriched (FDR-p = 0.0395) for transcripts from genes associated with both pirfenidone and nintedanib treatment. Among these, 4 genes (*BASPI*, *HSD17B6*, *CDH11*, and *TNS1*) associated with pirfenidone were identified as novel genes not found in known pathways containing TGFβ1 (Table 1). Additionally, 5 genes (*CLINT1*, *CADM1*, *MTDH*, *SYDE1*, and *MCTS1*) associated with nintedanib treatment have not previously been classified as involved with TGFβ1 signaling (Table 1). One gene, *MYDGF*, was highlighted as being associated with both pirfenidone and nintedanib and not a member of known TGFβ1 signaling pathways (Table 1).

Pathway analysis of novel fibrotic biomarkers in context with known pathobiology. The genes in the blue module annotated as associated with pirfenidone are enriched for genes involved in the epithelial mesenchymal transition (Hallmark Epithelial Mesenchymal Transition, FDR p-value = 1.1×10^{-9}), tissue morphogenesis (GO Tissue Morphogenesis, FDR p-value = 3.2×10^{-5}), and the inflammatory response (Hallmark Inflammatory Response, FDR p-value = 9.6×10^{-5}), among others (Table S9). The genes in the blue module annotated as associated with nintedanib were enriched for genes involved in the response to wounding (GO Response to Wounding, FDR p-value = 2.2×10^{-2}), apoptosis (Hallmark Apoptosis, FDR p-value = 9.2×10^{-3}), and platelet derived growth factor binding (GO Platelet Derived Growth Factor Binding, FDR p-value = 4.7×10^{-3}) (Table S9).

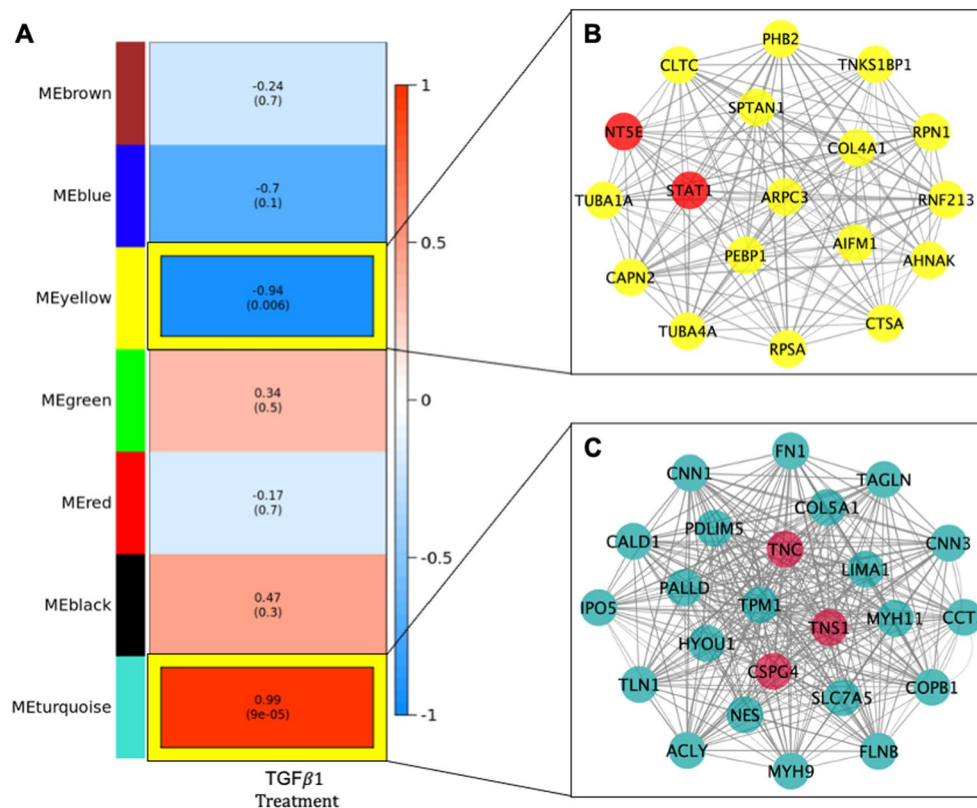


Figure 2. Weighted gene co-expression network analysis of proteomic data generated from IMR-90 cells with and without TGFβ treatment. **(A)** Proteomic Modules Associated with TGFβ1 Treatment in IMR-90 cells. Values represent correlation with p-values in parentheses between each module and trait. Heatmap shading corresponds to strength of association where darker red cells have higher upregulation and darker blue cells have higher downregulation based on correlation. Text outlined in yellow denotes result withstands Bonferroni correction for multiple testing based on the number of modules generated. **(B,C)** Network of hub proteins in proteomic modules significantly associated with TGFβ Treatment. Proteins with a kME larger than 0.90 were selected for visualization in the turquoise **(B)** and yellow **(C)** modules. The size of the circle in each network corresponds to increasing module membership and the thickness of the edge corresponds to increasing topological overlap (TOM), a measure of the strength of correlation between protein levels, which is the Pearson's correlation obtained from the adjacency matrix. Yellow nodes correspond to significant single proteins in the turquoise module associated with TGFβ1 treatment. Red nodes correspond to known targets of pirfenidone and/or nintedanib.

Gene name	Gene	Blue module membership
Pirfenidone only		
Brain abundant membrane attached signal protein 1	<i>BASP1</i>	0.76
Hydroxysteroid 17-beta dehydrogenase 6	<i>HSD17B6</i>	0.99
Cadherin 11	<i>CDH11</i>	0.92
Tensin 1	<i>TNS1</i>	0.93
Nintedanib only		
Clathrin interactor 1	<i>CLINT1</i>	0.97
Cell adhesion molecule 1	<i>CADM1</i>	0.95
Metadherin	<i>MTDH</i>	0.90
Synapse defective rho GTPase homolog 1	<i>SYDE1</i>	0.70
MCTS1 re-initiation and release factor	<i>MCTS1</i>	0.62
Pirfenidone and nintedanib		
Myeloid derived growth factor	<i>MYDGF</i>	0.94

Table 1. Novel genes targeting pirfenidone, nintedanib, or both pirfenidone and nintedanib. Novel is defined as not previously identified in known TGFβ1 signaling pathways.

Genes annotated as associated with both pirfenidone and nintedanib in the blue module were enriched with members of the collagen containing extracellular matrix (GO Collagen Containing Extracellular Matrix, FDR p -value = 1.8×10^{-12}), those involved in mesenchymal cell differentiation (GO Mesenchymal Cell Differentiation, FDR p -value = 3.2×10^{-9}), and lung fibrosis (WP Lung Fibrosis, 1.8×10^{-6}), among others (Table S9).

Evidence for drug repurposing based on known drug targets of novel TGF β 1 genes. Using the Clue Drug Repurposing Hub, succinic acid was highlighted as a metabolite regulated by the protein encoded by *HSD17B6* (Hydroxysteroid 17-Beta Dehydrogenase 6) (Table S10)¹⁹.

Discussion

In this study, we provide new insights into the anti-fibrotic mechanisms of pirfenidone and nintedanib and uncover novel drug targets in TGF β 1-driven fibrosis. Using an integrative 'omics approach, we identified modules of co-expressed transcripts and proteins induced by TGF β 1 in fibroblasts, key effector cells of tissue/organ fibrosis. We used an agnostic approach to identify markers of early TGF β 1-induced fibrogenesis and found TGF β 1 induced genes annotated as associated with pirfenidone and nintedanib treatment. In particular, we demonstrate evidence for a common mechanism of action for pirfenidone and nintedanib in modulating TGF β 1-induced *MYDGF*. Pirfenidone and nintedanib may mediate their anti-fibrotic effects via different mechanisms, such as *BASPI*, *HSD17B6*, *CDH11*, and *TNS1* with pirfenidone, and *CLINT1*, *CADMI*, *MTDH*, *SYDE1*, and *MCTS* with nintedanib.

Both pirfenidone and nintedanib likely influence the expression of several genes regulated by TGF β 1. Genes annotated as associated with pirfenidone treatment included *BASPI*, *HSD17B6*, *CDH11*, and *TNS1*. While *BASPI* and *HSD17B6* have been associated with multiple cancers, both *CDH11* and *TNS1* have been linked to mesenchymal activation and pulmonary fibrosis^{20–24}. The cadherin 11 protein, encoded by *CDH11*, is a member of a family of integral membrane proteins responsible for the mediation of calcium-dependent cell to cell adhesion and has been implicated in epithelial–mesenchymal transition and pulmonary fibrosis²³. Tensin 1, encoded by *TNS1*, is a key protein that is a component of fibrillar adhesions that attach to the extracellular matrix and is essential for myofibroblast differentiation²⁴. This may indicate the clinical efficacy of pirfenidone and nintedanib in patients with fibrotic lung disease may involve both shared (common) pathways and single genes that mediate pro-fibrotic effects. The multifunctional effects of these drugs may also explain the need for combinatorial therapeutic approaches or single agents with pleiotropic effects. We found *CLINT1*, *CADMI*, *MTDH*, *SYDE1*, and *MCTS1* are annotated as being associated with nintedanib treatment; of these, *CADMI* and *MTDH* have been implicated in fibrosis. *CADMI*, an immunoglobulin superfamily member, has been reported to contribute pro-fibrotic effects through direct effects on fibroblasts and indirect effects on mast cell adhesion^{25,26}. Metadherin (*MTDH*) has been shown to mediate changes consistent with epithelial–mesenchymal transition in kidney fibrosis²⁷.

In our study, *HSD17B6* (Hydroxysteroid 17-Beta Dehydrogenase 6) was identified as a member of the highly TGF β 1 correlated blue module and as a single gene significantly upregulated (log FC = 3.48, FDR p -value = 0.0021) in TGF β 1 treated cells. *HSD17B6* has oxidoreductase activity and plays a key role in androgen catabolism and previous studies have demonstrated an association between androgen deficiency and cavernosal fibrosis^{28,29}. In our in-silico analyses, *HSD17B6* was annotated as associated with pirfenidone. Previous studies have identified *GLRX*, also having oxidoreductase activity, as a therapeutic target of pirfenidone whose forced expression was sufficient to inhibit or reverse liver fibrosis³⁰. Interestingly, single cell RNA-seq demonstrated *HSD17B6* as highly expressed in mesothelial cells in IPF³¹. Mesothelial cells play a direct role in fibrogenesis, and antioxidants have been shown to alleviate TGF β 1 induced EMT in mesothelial cells^{32,33}. Using the Clue Drug Repurposing Hub, we were able to identify succinic acid as a likely target of the protein encoded by *HSD17B6*. Increased accumulation of succinate has been implicated as a promoter of the development of fibrosis in the lung and liver^{34,35}. Interestingly, succinic acid has been identified as an important signaling molecule in both pulmonary and liver fibrosis with the development of therapies targeting succinate proposed as a potential treatment to prevent and/or cure fibrosis in these tissues^{34,35}. Although succinate may have more affinity for its receptor (succinate receptor 1), its expression is varied according to tissue type³⁶. However, compared to other tissues in the body, including liver tissue, the expression of succinate receptor 1 in the lung is low³⁶. This suggests that succinate may mediate biological effects independent of its receptor(s) activation, such as post-translation modifications of proteins involving succinylation³⁷.

MYDGF (also known as IL-25), which encodes the Myeloid Derived Growth Factor protein, was co-expressed with genes up or down regulated in TGF β 1 induced model of fibrosis. To our knowledge, this is first time *MYDGF*, annotated as associated with pirfenidone and nintedanib treatment, has been shown to be regulated by TGF β 1³⁸. The biological function of *MYDGF* is relatively unknown; however, previous studies in mouse models of coronary artery disease have demonstrated that monocytes and macrophages secrete *MYDGF* as a protective and reparative response following myocardial infarction³⁹. *MYDGF* treatment was shown to reduce scar size and contractile dysfunction and has been identified as a potential therapeutic target for cardiac fibrosis^{39,40}. Although an agnostic approach was employed to identify *MYDGF* as a novel TGF β 1 signaling target, it was co-expressed with other known TGF β 1 signaling targets (*VEGF* and *ACTA2*, among others), providing additional support for this finding⁴¹.

Our study has several strengths and limitations. A strength of our study is that it presents the first integrated proteomic and transcriptomic network analysis of human lung fibroblasts treated with TGF β 1. We also took advantage of a network-based approach which highlighted new protein and gene expression modules associated with TGF β 1 treatment. Through the integration of significant proteomic modules into the transcriptomic module-trait correlation, we were able to systematically investigate the interplay of downstream targets associated with TGF β 1. Further, PSEA and GSEA indicated proteins and genes in our agnostically derived modules were

significantly enriched for pathways with known relevant to TGF β 1 biology. Among the limitations, we analyzed proteomic and transcriptomic data generated at a single time point. A longitudinal analysis would have provided a deeper understanding of the changes in protein levels and gene expression over time. However, by integrating our proteomic and transcriptomic results, we were able to observe the downstream mechanisms contributing to fibrogenesis. Further, we limited our studies to exploring primary effects of TGF β 1 signaling in fibroblasts obtained from a cell line as opposed to primary cells. Another limitation of this study is the lack of an equivalent study in which we could bioinformatically validate the genes annotated as associated with pirfenidone and nintedanib not found in known pathways containing TGF β 1. However, these findings have the potential to identify novel pathways and/or molecular targets within these pathways due to the agnostic approach employed by this study. Follow-up via replication is an important future direction for our understanding of the robustness of our novel finding. Further, unlike transcripts, proteins cannot be amplified, and as this was not a targeted study, but rather a global discovery proteomics analysis, we are limited to what is observed. However, with the use of systems analysis combined with transcriptomics, this data pinpoints key pathways and molecular changes with high confidence; therefore, the lack of observation of a few known proteins should not detract from the utility of this kind of multi-Omics study. Finally, the genes associated with pirfenidone and nintedanib obtained from Coremine may have data not just from IMR-90 cells, but also from other cell lines that may confound the results.

In conclusion, we identified potential molecular targets involved in TGF β 1 signaling driving myofibroblast differentiation and preclinical fibrosis development. We also demonstrated the utility of integrative network analyses to identify novel molecular targets and pathways that help elucidate the role of TGF β 1 signaling in the anti-fibrotic actions of pirfenidone and nintedanib. Further studies of these novel targets are warranted to confirm reproducibility and potential therapeutic efficacy in relevant pre-clinical models.

Methods

Study design and ethics. IRB approval was obtained from the University of Alabama at Birmingham for analyses presented in this manuscript. All statistical analyses, including WGCNA, were performed in R, version 3.6.0. Cytoscape, version 3.8.1, was used to visualize molecular pathways of interest.

Cell culture. Human fetal lung mesenchymal cells (hFLMCs; IMR-90 cells) were obtained from Coriell Cell Repositories, Institute for Medical Research, Camden, NJ. IMR-90 cells were cultured in DMEM (Life Technologies, Inc.) supplemented with 10% fetal calf serum (Hyclone Laboratories, Logan, UT), 100 U/ml penicillin, 100 μ g/ml streptomycin, and 1.25 μ g/ml amphotericin B. IMR-90 cells were incubated at 37 °C in 5% CO₂ and 95% air. For both transcriptomic and proteomic analyses, a total of 6 flasks of IMR-90 cells with (N = 3) replicates per group, TGF β treatment (2 ng/ml for 16 h) and no TGF β treatment, were included in this study.

Transcriptomic data. Gene expression data was obtained from GEO (GEO: GSE17518). Sample preparation and quality control for gene expression data have been previously described⁴². Briefly, gene expression data was generated from IMR-90 cells and profiled using the Affymetrix Human U133A array (Fig. S1). For (N = 3) samples treated with TGF β , mRNA was collected 48 h post treatment^{42,43}. Gene expression data was normalized using log transformation. Probes with low variance and those which did not annotate within a specific gene were removed, leaving a final sample size of 6,456 probes for analysis. The amount of missing data for each probe was assessed for quality control, and no probes were removed due to missingness, defined as missing in more than 1 of either control or treatment samples (Fig. S1).

Proteomic data. Proteomics analyses were carried out as previously referenced with minor changes (Ludwig et al, under section 2.5 nLC-ESI-MS2 under Protein IDs for GeLC)⁴⁴. The protein fractions were quantified, 40 μ g of protein per sample were reduced with DTT and denatured at 70 °C for 35 min prior to loading onto 10% Bis-Tris Protein gels and separated. The gels were stained overnight with colloidal Coomassie for visualization purposes, the entire gel lane was cut into 6-MW fractions, and each plug was equilibrated in 100 mM ammonium bicarbonate (AmBc), and digested overnight with Trypsin Gold, Mass Spectrometry Grade (Promega, Cat.# V5280) following manufacturer's instruction. Peptide extracts were reconstituted in 0.1% Formic Acid/ddH₂O at 0.1 μ g/ μ L. Mass spectrometry was carried out, and the data was processed, searched, filtered, grouped, and quantified, as previously reported in detail⁴⁵. Following protein identification, and relative quantification by normalized spectral counting (NSC), the most statistically significant changed proteins from each pairwise comparison were analyzed⁴⁵.

An overview of the computational integrated 'omics analysis methods used are provided in Fig. S1. Peptides missing in more than 1 of either control or treatment samples were removed from the analysis. Proteomic data was normalized using a log transformation and proteins with low variance were removed from analyses, leaving a final sample size of N = 533 peptides for analyses (Fig. S1). The R MICE package was used to impute missing data⁴⁶.

Single transcript analysis. Linear regression models were fit for each single-gene transcript using an empirical Bayes method to determine if any significant single-gene transcripts were associated with TGF β 1 treatment. False discovery rate (FDR) p-values were calculated based on the number of transcripts in the array (N = 22,284). In instances where multiple probes mapping to a single gene, the probe with the highest mean expression value was selected⁴⁷.

Single peptide analysis. Linear regression models were fit for each protein using an empirical Bayes method to determine if any significant proteins were associated with TGF β 1 treatment. False discovery rate (FDR) p-values were calculated based on the number of proteins identified from the column post-quality control (N = 533).

Weighted gene co-expression (WGCNA) analysis. WGCNA¹⁸ is an established network analysis method which maximizes the statistical power of complex analyses by taking into account the correlated nature of biological networks. WGCNA can be used to translate and integrate 'omics data into networks of co-expressed biomarkers which can be tested for association with phenotypes. Focusing on networks as opposed to single, candidate biomarkers provide a biologically relevant approach to visualizing 'omic pathways through the observation of combined influence and interrelation of multiple molecular layers on the disease process. In addition, WGCNA networks are generated using an agnostic approach as opposed to reliance on known biological pathways to identify a wider scope of novel biomarkers. The WGCNA R package¹⁸ was used to identify modules of co-expressed proteins and genes (also termed eigenprotein and eigengene), which consist of groups of proteins and genes with similar protein and gene expression patterns. A signed correlation network was built using a Pearson's correlation with a soft thresholding power of 6 for proteomic data and 18 for transcriptomic data. The soft thresholding power was determined using the criterion of approximate scale-free topology. Using hierarchical clustering, WGCNA partitions the total set of genes or proteins into distinct, non-overlapping modules labeled by color. In addition, each module corresponds to a module eigengene, which is the weighted average expression profile of the module. Each proteomic module generated by WGCNA was tested for correlation with TGF β 1 treatment. Transcriptomic modules generated by WGCNA were tested for correlation with TGF β 1 treatment and significant proteomic modules. Proteomic module-trait correlations with a Bonferroni corrected P-value less than 0.007 accounting for the number of modules generated by WGCNA (N = 7 modules) were considered statistically significant. Transcriptomic module-trait correlations with a Bonferroni corrected P-value less than 0.0045 accounting for the number of modules generated by WGCNA (N = 11 modules) were considered statistically significant.

Identification of genes associated with pirfenidone and nintedanib treatment. Genes annotated as associated with pirfenidone and nintedanib treatment were identified using Coremine medical (<https://coremine.com/medical/>)⁴⁸. Coremine presents results as a network that describes relationship between search terms (pirfenidone and nintedanib) and biological terms (including, but not limited to, gene and proteins terms) discovered through text-mining of the MEDLINE database (i.e. titles and abstracts contained in PubMed)⁴⁸. The strength of the association between search terms and biological terms is based on the number of co-occurrences of both terms in the literature.

Pathway analysis. We constructed both single gene expression probe/peptide models and network models. Protein set enrichment analysis (PSEA) of single proteins and gene set enrichment analysis (GSEA) of single genes significantly differentially associated with TGF β 1 treatment was performed using the Molecular Signatures Database (MSigDB) v7.0⁴⁹. This included gene set collections comprised of the hallmark gene set (N = 50 gene sets), GO gene sets (N = 9996 gene sets), and curated gene sets (N = 5501 gene sets). In addition, downstream analysis also took advantage of annotations described above. GSEA was also performed on the subset of proteins and genes that were identified as both significant single proteins and single genes significantly differentially associated with TGF β 1 treatment. Additionally, GSEA was performed on genes annotated as associated with pirfenidone only, nintedanib only, and by both pirfenidone and nintedanib within transcriptomic modules significantly correlated with proteomic modules. The Cytoscape EnrichmentMap R package⁵⁰ was used to visualize hub proteins in proteomic modules significantly correlated with TGF β 1 treatment and hub genes in transcriptomic modules significantly correlated with TGF β 1 treatment. Hub proteins were defined as proteins having a kME value greater than 0.90 and hub genes were defined as genes having a kME greater than 0.99. PSEA and GSEA significance was defined as having an FDR-q p-value less than 0.05.

Downstream interpretation and annotation of WGCNA module members. First, we compiled a list of pathways known to be involved with TGF β 1 signaling if they occurred MSigDB v7.0 along with TGF β 1 (see supplementary Table S7). In addition, drugs with the potential to be repurposed were identified using the Broad Institute CLUE drug repurposing tool (<https://clue.io/repurposing-app>)⁵¹. This information was used to annotate genes associated with pirfenidone only, nintedanib only, and by both pirfenidone and nintedanib within transcriptomic modules significantly correlated with proteomic modules and TGF β 1 treatment.

Data availability

The transcriptomics data analyzed in the current study (GSE17518) are available through the Gene Expression Omnibus (<https://www.ncbi.nlm.nih.gov/geo/>). The proteomics data analyzed in the current study are available from the corresponding author on reasonable request.

Received: 28 October 2021; Accepted: 14 February 2022

Published online: 23 February 2022

References

1. Wynn, T. A. Common and unique mechanisms regulate fibrosis in various fibroproliferative diseases. *J. Clin. Investig.* **117**(3), 524–529. <https://doi.org/10.1172/JCI31487> (2007).

2. Thannickal, V. J., Zhou, Y., Gaggar, A. & Duncan, S. R. Fibrosis: Ultimate and proximate causes. *J. Clin. Investig.* **124**(11), 4673–4677. <https://doi.org/10.1172/JCI74368> (2014).
3. Penke, L. R. & Peters-Golden, M. Molecular determinants of mesenchymal cell activation in fibroproliferative diseases. *Cell Mol. Life Sci.* **76**(21), 4179–4201. <https://doi.org/10.1007/s00018-019-03212-3> (2019).
4. Jin, J. *et al.* Pirfenidone attenuates lung fibrotic fibroblast responses to transforming growth factor- β 1. *Respir. Res.* **20**(1), 119. <https://doi.org/10.1186/s12931-019-1093-z> (2019).
5. Kim, K. K., Sheppard, D. & Chapman, H. A. TGF- β 1 signaling and tissue fibrosis. *Cold Spring Harb. Perspect. Biol.* **10**(4), 04. <https://doi.org/10.1101/cshperspect.a022293> (2018).
6. Saito, S., Alkhatib, A., Kolls, J. K., Kondoh, Y. & Lasky, J. A. Pharmacotherapy and adjunctive treatment for idiopathic pulmonary fibrosis (IPF). *J. Thorac. Dis.* **11**(Suppl 14), S1740–S1754. <https://doi.org/10.21037/jtd.2019.04.62> (2019).
7. Olson, A. L., Gifford, A. H., Inase, N., Fernández Pérez, E. R. & Suda, T. The epidemiology of idiopathic pulmonary fibrosis and interstitial lung diseases at risk of a progressive-fibrosing phenotype. *Eur. Respir. Rev.* <https://doi.org/10.1183/16000617.0077-2018> (2018).
8. Lederer, D. J. & Martinez, F. J. Idiopathic pulmonary fibrosis. *N. Engl. J. Med.* **378**(19), 1811–1823. <https://doi.org/10.1056/NEJMr1705751> (2018).
9. Gulati, S. & Thannickal, V. J. The aging lung and idiopathic pulmonary fibrosis. *Am. J. Med. Sci.* **357**(5), 384–389. <https://doi.org/10.1016/j.amjms.2019.02.008> (2019).
10. Blackwell, T. S. *et al.* Future directions in idiopathic pulmonary fibrosis research. An NHLBI workshop report. *Am. J. Respir. Crit. Care Med.* **189**(2), 214–222. <https://doi.org/10.1164/rccm.201306-1141WS> (2014).
11. Rangarajan, S., Locy, M. L., Luckhardt, T. R. & Thannickal, V. J. Targeted therapy for idiopathic pulmonary fibrosis: Where to now?. *Drugs* **76**(3), 291–300. <https://doi.org/10.1007/s40265-015-0523-6> (2016).
12. Richeldi, L. *et al.* Efficacy and safety of nintedanib in idiopathic pulmonary fibrosis. *N. Engl. J. Med.* **370**(22), 2071–2082. <https://doi.org/10.1056/NEJMoa1402584> (2014).
13. Gulati, S. & Luckhardt, T. R. Updated evaluation of the safety, efficacy and tolerability of pirfenidone in the treatment of idiopathic pulmonary fibrosis. *Drug Healthc. Patient Saf.* **12**, 85–94. <https://doi.org/10.2147/DHPS.S224007> (2020).
14. Varone, F., Sgalla, G., Iovene, B., Bruni, T. & Richeldi, L. Nintedanib for the treatment of idiopathic pulmonary fibrosis. *Expert Opin. Pharmacother.* **19**(2), 167–175. <https://doi.org/10.1080/14656566.2018.1425681> (2018).
15. Ballester, B., Milara, J. & Cortijo, J. Pirfenidone anti-fibrotic effects are partially mediated by the inhibition of MUC1 bioactivation. *Oncotarget* **11**(15), 1306–1320. <https://doi.org/10.18632/oncotarget.27526> (2020).
16. Kwapiszewska, G. *et al.* Transcriptome profiling reveals the complexity of pirfenidone effects in idiopathic pulmonary fibrosis. *Eur. Respir. J.* **52**(5), 11. <https://doi.org/10.1183/13993003.00564-2018> (2018).
17. Spagnolo, P. *et al.* Idiopathic pulmonary fibrosis: Disease mechanisms and drug development. *Pharmacol. Ther.* **222**, 107798. <https://doi.org/10.1016/j.pharmthera.2020.107798> (2021).
18. Langfelder, P. & Horvath, S. WGCNA: An R package for weighted correlation network analysis. *BMC Bioinform.* **9**, 559. <https://doi.org/10.1186/1471-2105-9-559> (2008).
19. Wishart, D. *et al.* T3DB: The toxic exposome database. *Nucleic Acids Res.* **43**(Database issue), D928–D934. <https://doi.org/10.1093/nar/gku1004> (2015).
20. Wang, H., Leinwand, L. A. & Anseth, K. S. Roles of transforming growth factor- β 1 and OB-cadherin in porcine cardiac valve myofibroblast differentiation. *FASEB J.* **28**(10), 4551–4562. <https://doi.org/10.1096/fj.14-254623> (2014).
21. Zhang, A., Zhang, J., Plymate, S. & Mostaghel, E. A. Classical and non-classical roles for pre-receptor control of DHT metabolism in prostate cancer progression. *Horm. Cancer* **7**(2), 104–113. <https://doi.org/10.1007/s12672-016-0250-9> (2016).
22. Hartl, M. & Schneider, R. A unique family of neuronal signaling proteins implicated in oncogenesis and tumor suppression. *Front. Oncol.* **9**, 289. <https://doi.org/10.3389/fonc.2019.00289> (2019).
23. Schneider, D. J. *et al.* Cadherin-11 contributes to pulmonary fibrosis: Potential role in TGF- β production and epithelial to mesenchymal transition. *FASEB J.* **26**(2), 503–512. <https://doi.org/10.1096/fj.11-186098> (2012).
24. Bernau, K. *et al.* Tensin 1 is essential for myofibroblast differentiation and extracellular matrix formation. *Am. J. Respir. Cell Mol. Biol.* **56**(4), 465–476. <https://doi.org/10.1165/rcmb.2016-0104OC> (2017).
25. Moiseeva, E. P., Roach, K. M., Leyland, M. L. & Bradding, P. CADM1 is a key receptor mediating human mast cell adhesion to human lung fibroblasts and airway smooth muscle cells. *PLoS ONE* **8**(4), e61579. <https://doi.org/10.1371/journal.pone.0061579> (2013).
26. Cao, W., Shi, P. & Ge, J. J. miR-21 enhances cardiac fibrotic remodeling and fibroblast proliferation via CADM1/STAT3 pathway. *BMC Cardiovasc. Disord.* **17**(1), 88. <https://doi.org/10.1186/s12872-017-0520-7> (2017).
27. Peng, F. *et al.* Michelolide ameliorates renal fibrosis by suppressing the Mtdh/BMP/MAKP pathway. *Lab. Investig.* **99**(8), 1092–1106. <https://doi.org/10.1038/s41374-019-0245-6> (2019).
28. Cho, M. C., Song, W. H. & Paick, J. S. Suppression of cavernosal fibrosis in a rat model. *Sex. Med. Rev.* **6**(4), 572–582. <https://doi.org/10.1016/j.sxmr.2018.02.007> (2018).
29. Moeller, G. & Adamski, J. Multifunctionality of human 17 β -hydroxysteroid dehydrogenases. *Mol. Cell Endocrinol.* **248**(1–2), 47–55. <https://doi.org/10.1016/j.mce.2005.11.031> (2006).
30. Xi, Y. *et al.* The anti-fibrotic drug pirfenidone inhibits liver fibrosis by targeting the small oxidoreductase glutaredoxin-1. *Sci. Adv.* **7**(36), eabg9241. <https://doi.org/10.1126/sciadv.abg9241> (2021).
31. Adams, T. S. *et al.* Single-cell RNA-seq reveals ectopic and aberrant lung-resident cell populations in idiopathic pulmonary fibrosis. *Sci. Adv.* **6**(28), eaba1983. <https://doi.org/10.1126/sciadv.aba1983> (2020).
32. Yu, M. *et al.* Effect of aldosterone on epithelial-to-mesenchymal transition of human peritoneal mesothelial cells. *Kidney Res. Clin. Pract.* **34**(2), 83–92. <https://doi.org/10.1016/j.krcp.2015.03.005> (2015).
33. Mutsaers, S. E. *et al.* Mesothelial cells in tissue repair and fibrosis. *Front. Pharmacol.* **6**, 113. <https://doi.org/10.3389/fphar.2015.00113> (2015).
34. Wang, Z. *et al.* Pharmaceutical targeting of succinate dehydrogenase in fibroblasts controls bleomycin-induced lung fibrosis. *Redox Biol.* **46**, 102082. <https://doi.org/10.1016/j.redox.2021.102082> (2021).
35. Cho, E. H. Succinate as a regulator of hepatic stellate cells in liver fibrosis. *Front. Endocrinol.* **9**, 455. <https://doi.org/10.3389/fendo.2018.00455> (2018).
36. Guo, Y., Cho, S. W., Saxena, D. & Li, X. Multifaceted actions of succinate as a signaling transmitter vary with its cellular locations. *Endocrinol. Metab.* **35**(1), 36–43. <https://doi.org/10.3803/EnM.2020.35.1.36> (2020).
37. Alleyn, M., Breitzig, M., Lockey, R. & Kolliputi, N. The dawn of succinylation: A posttranslational modification. *Am. J. Physiol. Cell Physiol.* **314**(2), C228–C232. <https://doi.org/10.1152/ajpcell.00148.2017> (2018).
38. Majewski, S. *et al.* Epithelial alarmins in serum and exhaled breath in patients with idiopathic pulmonary fibrosis: A prospective one-year follow-up cohort study. *J. Clin. Med.* **8**(10), 1590. <https://doi.org/10.3390/jcm8101590> (2019).
39. Korf-Klingebiel, M. *et al.* Myeloid-derived growth factor (C19orf10) mediates cardiac repair following myocardial infarction. *Nat. Med.* **21**(2), 140–149. <https://doi.org/10.1038/nm.3778> (2015).
40. Hinderer, S. & Schenke-Layland, K. Cardiac fibrosis—A short review of causes and therapeutic strategies. *Adv. Drug Deliv. Rev.* **146**, 77–82. <https://doi.org/10.1016/j.addr.2019.05.011> (2019).

41. Sano, M. *et al.* Potential role of transforming growth factor-beta 1/Smad signaling in secondary lymphedema after cancer surgery. *Cancer Sci.* **111**(7), 2620–2634. <https://doi.org/10.1111/cas.14457> (2020).
42. Hecker, L. *et al.* NADPH oxidase-4 mediates myofibroblast activation and fibrogenic responses to lung injury. *Nat. Med.* **15**(9), 1077–1081. <https://doi.org/10.1038/nm.2005> (2009).
43. Vittal, R. *et al.* Modulation of prosurvival signaling in fibroblasts by a protein kinase inhibitor protects against fibrotic tissue injury. *Am. J. Pathol.* **166**(2), 367–375. [https://doi.org/10.1016/S0002-9440\(10\)62260-2](https://doi.org/10.1016/S0002-9440(10)62260-2) (2005).
44. Zhou, J. J., Yip, W. K., Cho, M. H., Qiao, D., McDonald, M. L. & Laird N. M. A comparative analysis of family-based and population-based association tests using whole genome sequence data. In *BMC Proc.*, vol. 8, no. Suppl 1 Genetic Analysis Workshop 18 Vanessa Olmo, S33, (2014) <https://doi.org/10.1186/1753-6561-8-S1-S33>.
45. Ludwig, M. R. *et al.* Surveying the serologic proteome in a tissue-specific kras(G12D) knockin mouse model of pancreatic cancer. *Proteomics* **16**(3), 516–531. <https://doi.org/10.1002/pmic.201500133> (2016).
46. van Buuren, S. & Groothuis-Oudshoorn, K. mice: Multivariate imputation by chained equations in R. *J. Stat. Softw.* **45**(3), 67 (2011).
47. Miller, J. A. *et al.* Strategies for aggregating gene expression data: The collapseRows R function. *BMC Bioinform.* **12**, 322. <https://doi.org/10.1186/1471-2105-12-322> (2011).
48. de Leeuw, N. *et al.* Diagnostic interpretation of array data using public databases and internet sources. *Hum. Mutat.* **33**(6), 930–940. <https://doi.org/10.1002/humu.22049> (2012).
49. Liberzon, A. A description of the Molecular Signatures Database (MSigDB) web site. *Methods Mol. Biol.* **1150**, 153–160. https://doi.org/10.1007/978-1-4939-0512-6_9 (2014).
50. Shannon, P. *et al.* Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**(11), 2498–2504. <https://doi.org/10.1101/gr.1239303> (2003).
51. Corsello, S. M. *et al.* The Drug Repurposing Hub: A next-generation drug library and information resource. *Nat. Med.* **23**(4), 405–408. <https://doi.org/10.1038/nm.4306> (2017).

Acknowledgements

This manuscript was supported by the University of Alabama at Birmingham Predoctoral Training Program in Lung Diseases T32 HL134640, United States National Institutes of Health grants R01 HL067967 and R01 HL094230, and the UAB Comprehensive Cancer Center; Agency: National Institute of Health (NIH); Institute: National Cancer Institute (NCI), project # P30CA013148.

Author contributions

Planned the experiments V.J.T., J.A.M., and C.D. Carried out the experiments: S.A., C.D., J.A.M. Evaluated and interpreted the data: A.C.W., P.K.L., V.J.T., M.N.M., J.C., and E.R.N. Wrote and critically revised the manuscript: A.C.W., V.J.T., M.N.M., and E.R.N.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-07151-1>.

Correspondence and requests for materials should be addressed to M.-L.N.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

This is a U.S. Government work and not under copyright protection in the US; foreign copyright protection may apply 2022