Original Research

Molecular differences between stable idiopathic pulmonary fibrosis and its acute exacerbation

Junho Kang[1,†], Hye Ju Yeo[2,3,†], Yun Hak Kim[4,5,*], Woo Hyun Cho[2,3,*]

[1]*Medical Research Institute, Pusan National University, 46240 Busan, Republic of Korea,* [2]*Division of Allergy, Pulmonary and Critical Care Medicine, Department of Internal Medicine, Pusan National University Yangsan Hospital, 626-770 Yangsan, Republic of Korea,* [3]*Research Institute for Convergence of Biomedical Science and Technology, Pusan National University Yangsan Hospital, 626-770 Yangsan, Republic of Korea,* [4]*Department of Anatomy, School of Medicine, Pusan National University, 50612 Yangsan, Republic of Korea,* [5]*Department of Biomedical Informatics, School of Medicine, Pusan National University, 50612 Yangsan, Republic of Korea*

TABLE OF CONTENTS

1. Abstract

Introduction: The molecular mechanisms underlying acute exacerbations (AEs) of idiopathic pulmonary fibrosis (IPF) are poorly understood. To understand the gene expression patterns of the AEs of IPF, we studied gene expression profiling of AEs of IPF. **Methods**: The GEO datasets included in this study are GSE44723 and GSE10667, and in-house RNA-seq data were used. DEG analysis used the limma package, and the STRING database was used to construct the protein-protein interaction (PPI) network, and its functional role was investigated through gene ontology analysis. **Results**: The results of DEG analysis indicated 76 upregulated and 135 downregulated genes associated with an AE of IPF compared to stable IPF. The PPI network included three core modules containing 24 of the 211 DEGs. Eleven upregulated and six downregulated genes were evident in AEs of IPF compared with stable IPF after validation. The upregulated genes were associated with cell division. The downregulated genes were related to skeletal muscle differentiation and development. **Conclusion**: In previous studies, 17 genes were strongly associated with cell proliferation in various cell types. In particular, cyclin A2 (CCNA2) was overexpressed in the alveolar epithelium of the lungs presenting AEs of IPF. Aside from the previously described CCNA2, this study reveals 16 genes associated with AEs of IPF. This data could indicate new therapeutic targets and potential biomarkers for the AEs of IPF.

Submitted: 13 July 2021 Revised: 6 October 2021 Accepted: 25 October 2021 Published: 30 December 2021

2. Introduction

Idiopathic pulmonary fibrosis (IPF) is a progressive fibrotic lung disease with no clear etiology [1]. Currently, antifibrotic drugs constitute the main pharmacotherapy, with robust evidence from clinical studies [2]. Pirfenidone and nintedanib reportedly modestly reduce annual lung function decline. However, the specific molecular targets of these agents in the pathogenesis of IPF are not fully understood. Additionally, IPF presents with a wide spectrum of clinical phenotypes, especially during disease progression. The underlying molecular background of the various phenotypes is unknown, and molecular biomarkers of disease progression are lacking.

Among the wide variety of individual clinical courses, rapid acute decline in lung function, known as acute exacerbation (AE), occurs annually in many patients [3]. The optimal therapy for the AE of IPF has not been established. Although randomized trial data are lacking, treatment commonly involves systemic glucocorticoids [4]. Although lung transplantation provides a chance of survival, for those with acute deterioration, the perioperative risks of transplantation are clearly greater. For individuals transplanted in the context of mechanical ventilation or extracorporeal membrane oxygenation, short- and long-term survival after lung transplantation is decreased compared with that of stable patients undergoing lung transplantation [5]. Unfortunately, without transplantation, the median survival following an AE of IPF is only three to four months [6, 7].

Several previous studies identified the genes and/or pathways of IPF that are differentially expressed in comparison with the controls; thus, providing molecular signatures for IPF [8–12]. Furthermore, the peripheral blood profiling of IPF has suggested potential biomarkers. While these studies indicate pathways that could contribute to early, stable, or progressive IPF, knowledge of the pathways and mechanisms that contribute to the AEs of IPF remains limited. Few studies have screened differentially expressed genes (DEGs) associated with an AE of IPF using bioinformatics methods [13, 14]. Identifying the differences between steadily progressing IPF and the AEs of IPF can help in understanding the nature of the AEs. This knowledge would likely be valuable in the development of treatments for the AEs of IPF.

The present study investigated DEGs between cases of stable IPF and AE of IPF through the RNA sequencing (RNA-seq) of lung tissue and determined the hub genes via functional assays.

3. Materials and methods

3.1 Study population

Explanted lung tissue samples for RNA sequencing analysis were obtained from the Pusan National University Yangsan Hospital (PNUYH) biobank. These included lung tissues from three patients with IPF and lung tissues from three patients with AEs of IPF. The criteria for the diagnosis of IPF were those of the American Thoracic Society and European Respiratory Society [14]. An AE of IPF was defined as an acute, clinically significant respiratory deterioration characterized by evidence of new widespread alveolar abnormality according to the 2016 revised criteria [4]. All cases were reviewed by expert pulmonologists and pathologists. Detailed clinical information regarding the subjects is provided in Table 1. The mean forced vital capacity (expressed as a percentage of the normal expected value; FVC %) and diffusing capacity for carbon monoxide (expressed as a percentage of the normal expected value; DLCO %) of patients are provided in Table 1. Lung tissue acquisition and analysis were conducted with the appropriate approval from the Institutional Review Board of Pusan National University Yangsan Hospital (PNUYH IRB No 05-2020-019).

Table 1. Characteristics of patients with stable idiopathic pulmonary fibrosis and patients with acute exacerbation of IPF.

Variable	Stable IPF (n = 3)	AE of IPF (n = 3)
Age	64.3 ± 4.9	59.0 ± 10.0
Male	3 (100)	3 (100)
FVC %	58.7 ± 13.9	42 ± 0*
DLCO %	47.3 ± 33.6	34.0 ± 0*

Data presented as mean ± standard deviation or number (%). IPF, idiopathic pulmonary fibrosis; FVC, forced vital capacity; DLCO, diffusing capacity for carbon monoxide.
*denotes the last before AE of IPF.

3.2 RNA extraction, library construction, and sequencing

Total RNA from cell samples was extracted using a TRIzol reagent kit (Invitrogen), according to the manufacturer's protocol, and RNA integrity was assessed using a TapStation RNA screentape. After total RNA extraction, RNA libraries were independently prepared using the Illumina TruSeq Stranded Total RNA Library Prep Gold Kit (Illumina, Inc., San Diego, CA, USA). The cleaved RNA fragments were copied into first-strand cDNA using SuperScript II reverse transcriptase (Invitrogen, Carlsbad, CA, USA) using random primers. Qualified libraries were sequenced on an Illumina NovaSeq platform (Illumina, Inc., San Diego, CA, USA). Total RNA concentration was calculated using the Quant-IT RiboGreen Assay Kit (Invitrogen, Waltham, MA, USA). To determine the values of DV200 (percentage of RNA fragments >200 bp), samples were run on the TapeStation RNA screentape (Agilent, Wilmington, DE, USA). A total of 100 ng of total RNA was subjected to sequencing library construction using the Agilent SureSelect RNA Direct kit (Agilent, Wilmington, DE, USA) according to the manufacturer's protocol. Briefly, the to-

tal RNA was first fragmented into small pieces using divalent cations at elevated temperatures. The cleaved RNA fragments were copied into first-strand cDNA using random primers. This was followed by second-strand cDNA synthesis. These cDNA fragments then underwent end repair, the addition of a single "A" base, and the ligation of the adapters. The products were purified and enriched by PCR to create a cDNA library. To capture the human exonic region, the Agilent SureSelect XT Human All Exon v6+UTRs Kit (Agilent, Wilmington, DE, USA) was used according to the standard Agilent SureSelect Target Enrichment protocol. A cDNA library (25 ng) was mixed with hybridization buffers, blocking mixes, RNase block, and 5 μL of SureSelect XT Human All Exon v6+UTRs capture library (Agilent, Wilmington, DE, USA). Hybridization to the capture baits was conducted at 65 °C using the heated thermal cycler lid option at 105 °C for 24 h in a PCR machine. The captured library was washed and subjected to a second round of PCR amplification. The final purified product was then quantified using qPCR according to the qPCR Quantification Protocol Guide (KAPA Library Quantification kits for Illumina Sequencing platforms) and qualified using the TapeStation DNA screentape D1000 (Agilent, Wilmington, DE, USA). The indexed libraries were analyzed using the NovaSeq system (Daejeon, Korea). Paired-end (2 × 100 bp) sequencing was performed by Macrogen Inc. The gene expression data obtained through RNA-sequencing is provided in the **Supplementary file 1**.

3.3 Data acquisition

Expression profiling array data for IPF was obtained from the Gene Expression Omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/geo/; accession numbers GSE44723 and GSE10667). All datasets included data related to gene expression in the lung tissue of the AE group and/or that of the stable IPF group. The data were downloaded for DEG analysis. The characteristics of the included datasets are listed in Table 2 (Ref. [15, 16]).

3.4 Data preprocessing

DEGs of each dataset were identified using the GEO2R online analytic tool provided by GEO. Adjusted $p < 0.05$ and $|\log2FC| > 0$ were the cutoff criteria. For the adjusted p-value, false discovery rate methods were used. The probe identifier was converted to a genetic symbol based on the annotation file downloaded from the GEO. When multiple probes corresponded with a single gene, the average level of the probes was calculated as the expression value of the gene. Probes that did not correspond to the genetic symbols were removed. PNUYH RNA-seq data were quality checked using the fastqc (v 0.11.9) tool [17]. This check was performed after removing the sequence with a thread score <30 and Illumina adaptor sequence. The reference genome used for alignment was GRCH38 using the hisat2 tool (v 2.2.1) [18]. Stringtie (v 2.1.5) was used to obtain the read count for the DEG analysis [19]. DEGs were confirmed using the limma R package. Data preprocessing was performed using a voom provided in the limma package [20]. To screen the intersectional genes that were significantly expressed in each dataset, the R package "venn" was used to plot Venn diagrams [21]. Among the DEGs from the three datasets, only DEGs that commonly appeared in all three datasets were considered as significant DEGs and were selected for the subsequent analysis.

3.5 Protein–protein interaction (PPI) network construction and module selection

A PPI network was constructed using the Search Tool for the Retrieval of Interacting Genes (STRING) database [22] to identify the relevant pathways and functions of common DEGs. The minimum required interaction score used to construct the PPI network was 0.4. The molecular complex detection (MCODE) [23] was used for optimal module selection. The parameters used in MCODE were: degree cutoff = 2; cluster finding, node score cutoff = 0.2; k-core = 2; and maximum depth = 100. CytoHubba [24] was used to identify the highest-linkage hub genes in the network. The parameters used in CytoHubba were the top 10 nodes ranked by degree and the display option in which the expanded subnetwork was displayed.

3.6 Functional analysis [gene ontology (GO) analysis]

GO analyses were performed using the R package "clusterProfiler" (v 3.14.3) [25]. The common DEGs identified in each dataset were included in the GO term enrichment analysis, which was performed separately for up- and downregulated genes. Terms with five or more associated genes and two or more DEGs in each experiment were included. The adjusted p-values were calculated using the Benjamini-Hochberg method. Statistical significance was set at $p < 0.05$.

4. Results

4.1 Identification of common DEGs

The flowchart for this study is shown in Fig. 1. Detailed information on the datasets is presented in Table 1. DEG analysis of GSE10667 identified 4155 DEGs, of which 1832 genes were upregulated and 2325 genes were downregulated (Fig. 2). DEG analysis of PNUH RNA-seq identified 1819 DEGs, of which 897 genes were upregulated and 922 genes were downregulated (Fig. 2). DEG analysis results for each dataset are in **Supplementary files 2.1**, **2.2**, and **2.3**. We performed GO analysis of common DEGs and the results are in **Supplementary files 2.4** and **2.5**. Overlapping genes in each dataset included 76 upregulated genes and 135 downregulated genes, and the included genes list is in **Supplementary files 2.6**.

Table 2. Characteristics of the datasets.

	GSE10667		GSE44723		PNUYH	
	Stable	Acute exacerbation	Stable	Acute exacerbation	Stable	Acute exacerbation
Type of tissue	Lung tissue		Lung tissue		Lung tissue	
Platform	AGILENT-014850 Whole Human Genome Microarray 4 × 44K G4112F		Affymetrix Human Genome U133 Plus 2.0 Array		Illumina NovaSeq 6000	
Sample size	23	8	6	4	3	3
References	Konishi K *et al*., 2009 [15]		Sridhar S *et al*., 2013 [16]			

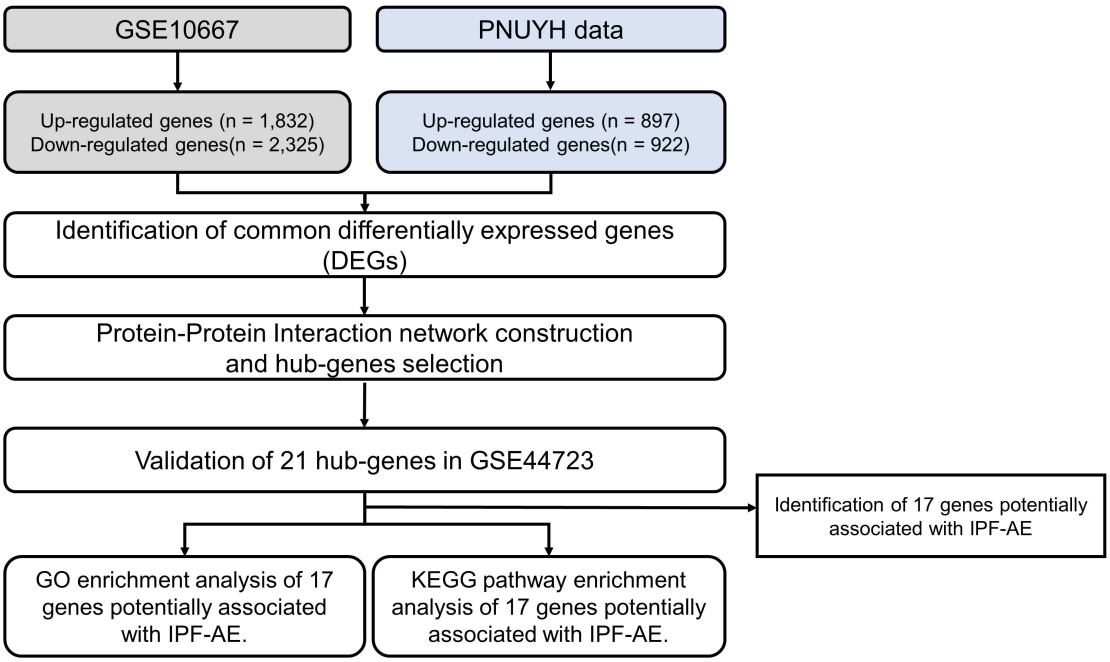

Fig. 1. Flow chart of the study.

4.2 PPI network construction and module selection

To identify the PPIs of common DEGs, the STRING network-based protein interaction assay was used to create a PPI network. To confirm the interaction between common DEGs, a PPI network was constructed. Three significant modules (confidence score >7) containing 24 out of 211 genes were identified (Fig. 3). The largest module comprising the upregulated genes consisted of *CCNA2*, *PTTG1*, *SKA3*, *SPAG5*, *TROAP*, *KIF14*, *ASPM*, *CENPE*, and *E2F7* (Fig. 3C). The second module comprising the upregulated genes consisted of histone-related genes (*HIST1H2AG*, *HIST3H2A*, *HIST1H2BJ*, *HIST1H2BI*, *HIST1H2BE*, and *HIST1H3B*; Fig. 3D). The third module comprising the downregulated genes contained six genes (*EGR1*, *NR4A1*, *FOSB*, *FOS*, *CYR61*, and *BTG2*; Fig. 3E).

4.3 Validation of the genes included in significant modules through GSE44723 expression data.

To validate the genes included in significant modules, the independent cohort GSE44723 was used. Eleven upregulated genes (*ASPM*, *CCNA2*, *CDC25B*, *CENPE*, *CLSPN*, *KIF14*, *PTTG1*, *RECQL4*, *SKA3*, *SPAG5*, and *TROAP*) and six downregulated genes (*BTG2*, *CYR61*, *EGR1*, *FOS*, *FOSB*, and *NR4A1*) were significantly altered in the cases presenting AEs of IPF compared to those presenting stable IPF (Fig. 4). Interestingly, the histone-associated genes and E2F7 gene were not significantly increased in AEs of IPF patients.

4.4 Functional analysis of the 17 genes in GSE44723

To investigate the functional role and the associated signaling pathways of the common DEGs, GO analysis was performed. The total significantly enriched GO terms from the up- and downregulated common DEGs are shown in Fig. 5. The top biological processes (BPs) of the eleven upregulated and six downregulated genes associated with the AEs of IPF were nuclear division, organelle fission, skeletal muscle cell differentiation, skeletal muscle tissue development, skeletal muscle organ development, response to mechanical stimulus, regulation of neuron death, and neuron death, respectively. The top cellular components of the upregulated and downregulated genes were spindle and transcription regulator complexes. The top molecular functions were microtubule binding, tubulin binding,

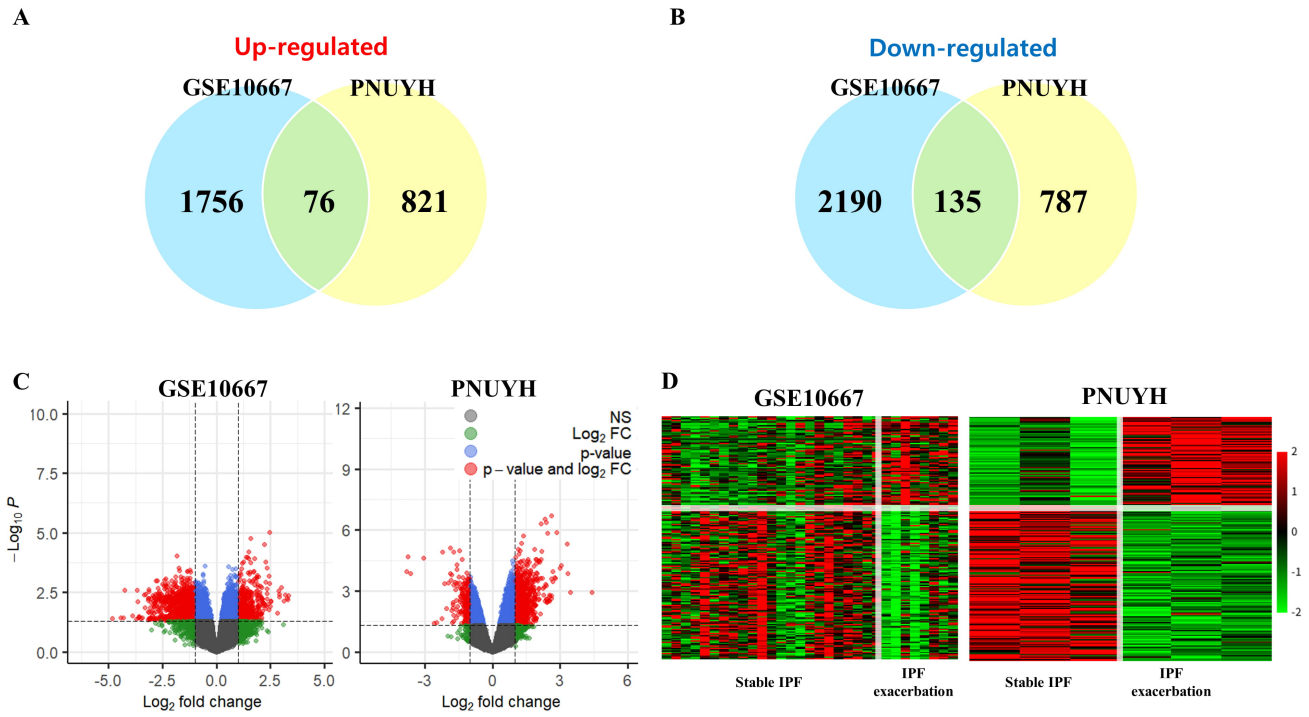

Fig. 2. Analysis of the differentially expressed genes. (A) Venn diagram of upregulated genes. (B) Venn diagram of downregulated genes. (C) Volcano plot of each cohort. In the plot, genes that met the *p*-value and $\log_2$ fold change cutoff value are denoted in red; genes that passed only the *p*-value are denoted in blue; a gene that passed only the $\log_2$ fold change is denoted in green; and genes that passed both the *p*-value and $\log_2$ fold change are denoted in black. (D) Heatmap of common DEGs for each cohort. The common upregulated genes are denoted in red and the common downregulated genes are denoted in green.

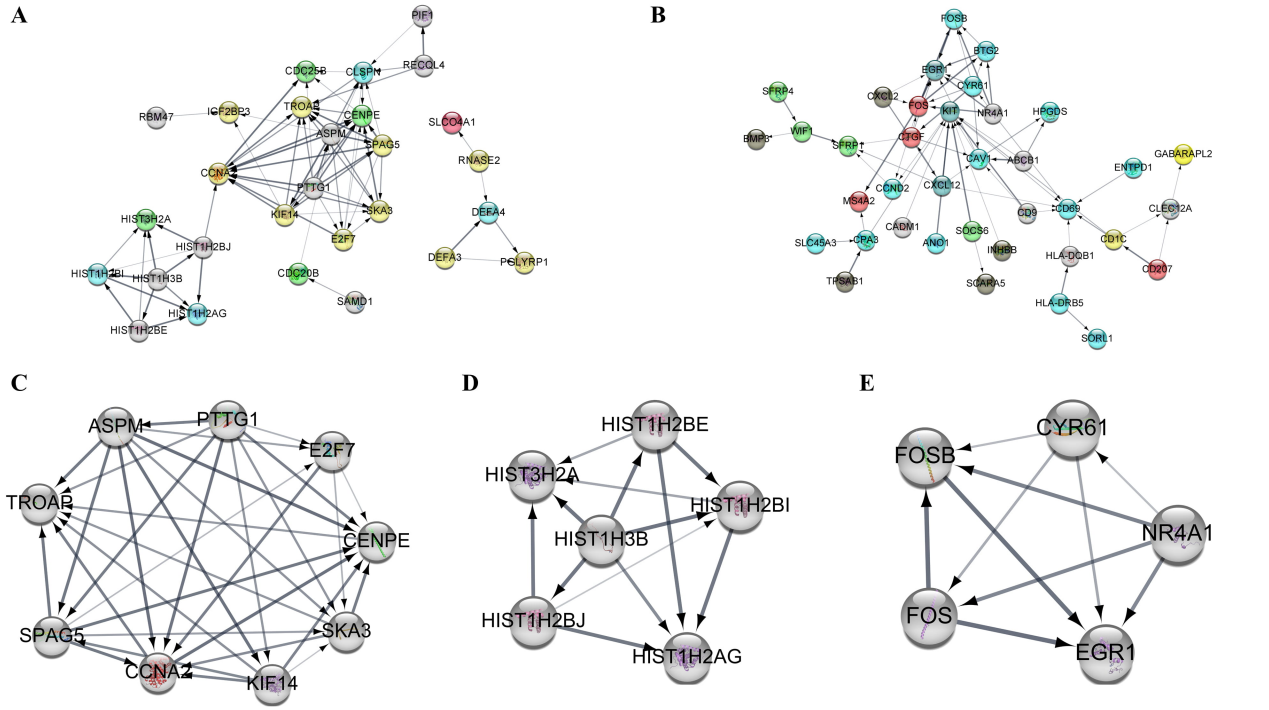

Fig. 3. Protein-protein interaction (PPI) network construction and significant module selection involving the common differentially expressed genes (DEGs). (A) PPI network of upregulated genes. (B) PPI network of downregulated genes. (C) The module with the highest confidence score among the PPI network of upregulated genes. (D) Module with the second-highest confidence score among the PPI network of upregulated genes. (E) The module with the highest confidence score among the PPI network of downregulated genes.

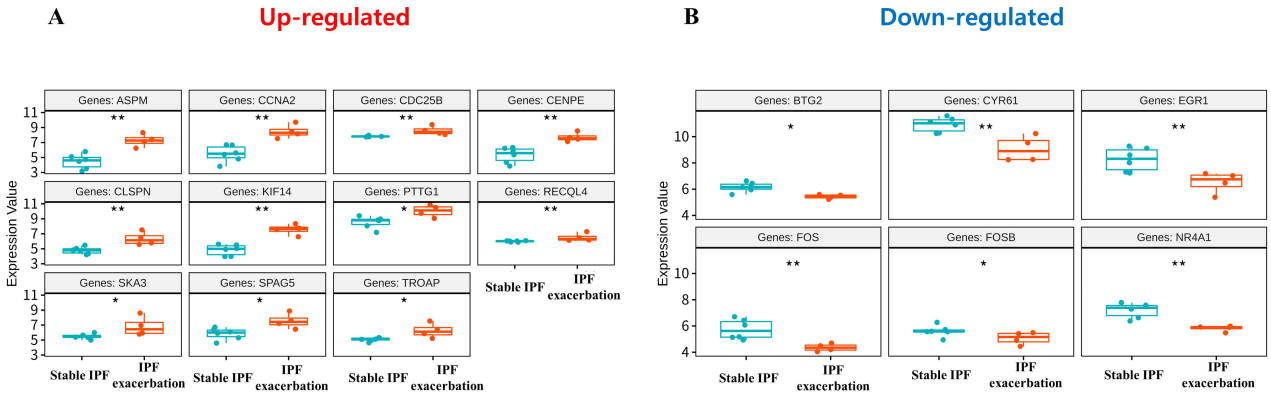

Fig. 4. Boxplots of 17 genes with statistically significant differences in expression values in GSE44723 among 24 genes included in significant modules. (A) Upregulated genes. (B) Downregulated genes. The blue boxplot indicates stable IPF and the red boxplot indicates acute exacerbation of IPF.

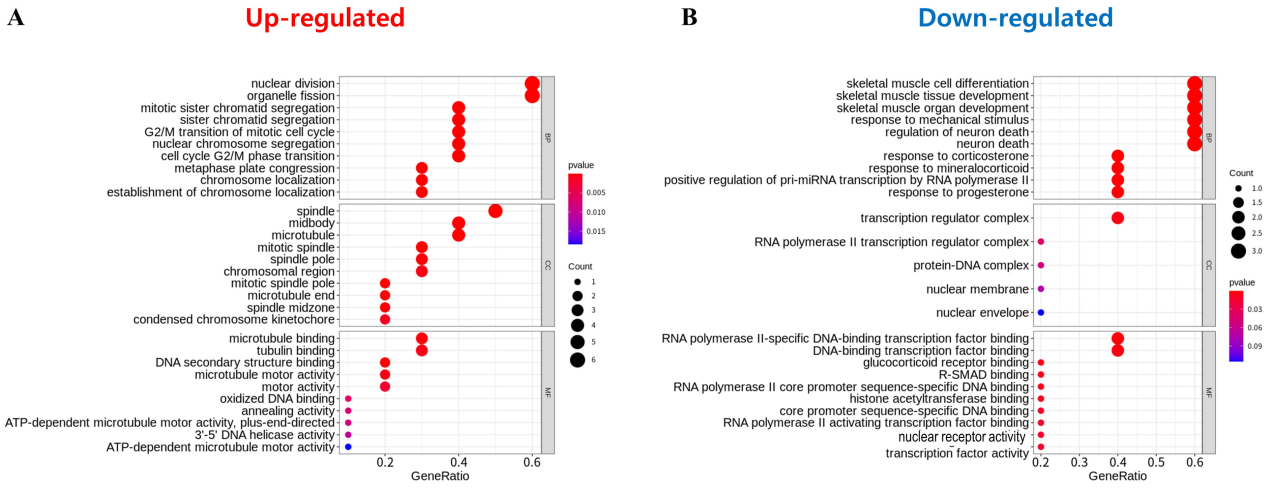

Fig. 5. Gene ontology assay of 17 genes identified through protein–protein interaction network and module selection. (A) Upregulated genes. (B) Downregulated genes. The gene ontology analysis result is composed of biological process (BP), cellular component (CC), and molecular function (MF) terms. The top 10 of each term is shown. The size of the circle indicates the number of genes corresponding to each term. The closer to red, the lower the *p*-value.

and RNA polymerase II-specific DNA binding transcription factor binding for upregulated proteins, and DNA binding transcription factor binding for downregulated proteins.

5. Discussion

The AE of IPF is a fatal disease phenotype, for which, there is presently no effective treatment. Annually, 10% to 15% of IPF cases have a potential risk of AE and the fatality rate exceeds approximately 90% [7]. To understand this refractory disease, genome-wide transcriptional profiling of different phenotypes of IPF may be carried out, which would further provide knowledge that could lead to an effective therapy for the AE of IPF.

Here, we selected cohorts from the GEO database (GSE10667 and GSE44723) to identify and validate the distinguishing molecular patterns associated with the AEs of IPF. The DEGs identified from the microarray data may be incorrect for reasons that include a lack of reproducibility and lack of prediction accuracy because of the small sample size [26]. To avoid possible problems, differential expression analysis was performed using the GSE10667 dataset and PNUYH RNA-seq data. The GSE44723 cohort was used to verify gene expression. Seventeen genes were potentially related to AEs of IPF after validation. The functional and signaling pathways of these genes were analyzed. In GO analysis, the upregulated genes were significantly enriched in cell division, such as mitotic sister chromatid segregation, nuclear division, and the G2/M transition of mitotic cell cycle. Several studies have demonstrated the respective upregulated genes to promote cell cycle progression and contribute to cell proliferation. SKA3 promotes cell proliferation and activates the phosphoinositide 3-kinase/Akt signaling pathway in cervical cancer [27]. Centrosome-associated protein E promotes cell prolifera-

tion in lung adenocarcinoma [28]. Kinesin family member 14 promotes cell proliferation via the activation of Akt [29]. *In vitro*, knockdown of trophinin associated protein significantly inhibits cell proliferation and the ability to transition from G1 to S phase [30]. Pituitary tumor transforming gene 1 overexpression enhances proliferation and upregulates the expression of cyclin B1, cyclin-dependent kinase 1, and c-Myc at both the protein and mRNA levels [31]. Claspin reportedly displayed a significantly higher relative increase in its levels than Ki67 in tumor samples compared with normal tissues, and was suggested as a marker of abnormal proliferation [32]. RecQ like helicase 4 and Sperm associated antigen 5 are essential for cell proliferation, cell cycle progression, and mitotic stability in human cells and promote DNA repair [33, 34]. Knockdown of cyclin A2 (CCNA2) inhibits colorectal cancer cell growth by impairing cell cycle progression and inducing apoptosis [35]. In particular, a prior study described that CCNA2 was overexpressed during AEs of IPF, unlike other genes, and appeared locally in the alveolar epithelium, but not in fibroblasts or myofibroblasts [14]. CCNA2 upregulation in AEs of IPF was confirmed in the present study. Eleven upregulated gene expression signatures, including CCNA2, could be potential biomarkers for evaluating patients with IPF.

The downregulated genes associated with AEs of IPF were *FOS, BTG2, FOSB, CYR61*, and *NR4A1*. GO analysis confirmed that they were most significantly enriched in skeletal muscle cell differentiation, development, and regulation of neuronal death. All downregulation was confirmed in previous studies to directly or indirectly affect cell proliferation when tumor suppressors are downregulated. Cysteine-rich angiogenic inducer 61 and nuclear receptor subfamily 4 group A member 1 inhibit cell proliferation as tumor suppressors in non-small cell lung cancer cells and triple-negative breast cancer, respectively [36, 37]. FOS and FOSB, also known together as "G0/G1 switch regulatory protein 3" in the FOS family, have been implicated as regulators of cell proliferation, differentiation, and transformation [38, 39]. BTG anti-proliferation factor 1 (BTG1) peaks in the G0/G1 phase of the cell cycle and decreases dramatically during the G1/S phase transition. This indicates that the overexpression of BTG1 normally inhibits cell growth [40]. Consistent with the results of the upregulated genes, the downregulated genes may affect cell proliferation.

6. Conclusions

Recently, many studies have addressed genomic alterations in specific diseases. Accordingly, tissue or blood sample RNA-seq, or single-cell RNA-seq data are providing a new paradigm for IPF research. However, RNA-seq data for AEs of IPF face some limitations. One of the main limitations is that it is very difficult to obtain tissue samples for tissue RNA-seq. In general, patients with AEs are critically ill, and tissue biopsy is not indicated, since it would further imperil the patients. In this study, we obtained whole Pleuroparenchymal fibroelastosis samples from explanted lungs after lung transplantation. This is a very rare and precious material for basic research on AEs of IPF. Because there were a limited number of samples, a pooled analysis was performed along with pre-existing dataset from other studies on AEs of IPF. To our knowledge, this is the first report of 17 genes involved in the AEs of IPF. The involvement of CCNA2 has been described previously. The remaining 16 genes described in this study are novel. It is essential to verify these genes through additional experiments. Pending these results, these genes may prove to be new therapeutic targets and potential predictive biomarkers for AEs of IPF.

7. Author contributions

YHJ, KYH, and CWH contributed to the conception of the work and critical revision of the manuscript. YHJ and CWH contributed to the data acquisition. KJ contributed to the interpretation of the data and analysis. KJ, YHJ, KYH, and CWH drafted and finalized the manuscript. All authors approved the submitted manuscript.

8. Ethics approval and consent to participate

The study was approved by the appropriate approval from the Institutional Review Board (PNUYH IRB No 05-2020-019).

9. Acknowledgment

Not applicable.

10. Funding

This study was supported by the Research Institute for Convergence of Biomedical Science and Technology Grant (30-2019-004), Pusan National University Yangsan Hospital. This work was supported by (NRF-2020R1F1A1075212, NRF-2020R1C1C1003741, NRF-2018R1A5A2023879).

11. Conflict of interest

The authors declare no conflict of interest.

12. References

[1] Raghu G, Collard HR, Egan JJ, Martinez FJ, Behr J, Brown KK, *et al*. An Official ATS/ERS/JRS/ALAT Statement: Idiopathic Pulmonary Fibrosis: Evidence-based Guidelines for Diagnosis and Management. American Journal of Respiratory and Critical Care Medicine. 2011; 183: 788–824.

[2] Azuma A. Pirfenidone: antifibrotic agent for idiopathic pulmonary fibrosis. Expert Review of Respiratory Medicine. 2010; 4: 301–310.

[3] Kim DS. Acute exacerbation of idiopathic pulmonary fibrosis: frequency and clinical features. European Respiratory Journal. 2006; 27: 143–150.

[4] Collard HR, Ryerson CJ, Corte TJ, Jenkins G, Kondoh Y, Lederer DJ, *et al*. Acute Exacerbation of Idiopathic Pulmonary Fibrosis. an International Working Group Report. American Journal of Respiratory and Critical Care Medicine. 2016; 194: 265–275.

[5] Yusen RD, Edwards LB, Kucheryavaya AY, Benden C, Dipchand AI, Goldfarb SB, *et al*. The Registry of the International Society for Heart and Lung Transplantation: Thirty-second Official Adult Lung and Heart-Lung Transplantation Report—2015; Focus Theme: Early Graft Failure. The Journal of Heart and Lung Transplantation. 2015; 34: 1264–1277.

[6] Travis WD, Costabel U, Hansell DM, King TE, Lynch DA, Nicholson AG, *et al*. An Official American Thoracic Society/European Respiratory Society Statement: Update of the International Multidisciplinary Classification of the Idiopathic Interstitial Pneumonias. American Journal of Respiratory and Critical Care Medicine. 2013; 188: 733–748.

[7] Juarez MM, Chan AL, Norris AG, Morrissey BM, Albertson TE. Acute exacerbation of idiopathic pulmonary fibrosis-a review of current and novel pharmacotherapies. Journal of Thoracic Disease. 2015; 7: 499–519.

[8] Yao C, Guan X, Carraro G, Parimon T, Liu X, Huang G, *et al*. Senescence of Alveolar Type 2 Cells Drives Progressive Pulmonary Fibrosis. American Journal of Respiratory and Critical Care Medicine. 2021; 203: 707–717.

[9] Sivakumar P, Thompson JR, Ammar R, Porteous M, McCoubrey C, Cantu E, *et al*. RNA sequencing of transplant-stage idiopathic pulmonary fibrosis lung reveals unique pathway regulation. ERJ Open Research. 2019; 5: 00117-2019.

[10] Xu Y, Mizuno T, Sridharan A, Du Y, Guo M, Tang J, *et al*. Single-cell RNA sequencing identifies diverse roles of epithelial cells in idiopathic pulmonary fibrosis. JCI Insight. 2016; 1: e90558.

[11] Vukmirovic M, Herazo-Maya JD, Blackmon J, Skodric-Trifunovic V, Jovanovic D, Pavlovic S, *et al*. Identification and validation of differentially expressed transcripts by RNA-sequencing of formalin-fixed, paraffin-embedded (FFPE) lung tissue from patients with Idiopathic Pulmonary Fibrosis. BMC Pulmonary Medicine. 2017; 17: 15.

[12] Nance T, Smith KS, Anaya V, Richardson R, Ho L, Pala M, *et al*. Transcriptome analysis reveals differential splicing events in IPF lung tissue. PLoS ONE. 2014; 9: e97550.

[13] Carleo A, Landi C, Prasse A, Bergantini L, D'Alessandro M, Cameli P, *et al*. Proteomic characterization of idiopathic pulmonary fibrosis patients: stable versus acute exacerbation. Monaldi Archives for Chest Disease. 2020; 90.

[14] Raghu G, Remy-Jardin M, Myers JL, Richeldi L, Ryerson CJ, Lederer DJ, *et al*. Diagnosis of Idiopathic Pulmonary Fibrosis. An Official ATS/ERS/JRS/ALAT Clinical Practice Guideline. American Journal of Respiratory and Critical Care Medicine. 2018; 198: e44–e68.

[15] Konishi K, Gibson KF, Lindell KO, Richards TJ, Zhang Y, Dhir R, *et al*. Gene Expression Profiles of Acute Exacerbations of Idiopathic Pulmonary Fibrosis. American Journal of Respiratory and Critical Care Medicine. 2009; 180: 167–175.

[16] Peng R, Sridhar S, Tyagi G, Phillips JE, Garrido R, Harris P, *et al*. Bleomycin induces molecular changes directly relevant to idiopathic pulmonary fibrosis: a model for "active" disease. PLoS ONE. 2013; 8: e59348.

[17] Andrews S. FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]. 2015. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/ (Accessed: 1 August 2021).

[18] Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nature Biotechnology. 2019; 37: 907–915.

[19] Pertea M, Pertea GM, Antonescu CM, Chang T, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nature Biotechnology. 2015; 33: 290–295.

[20] Smyth GK, Ritchie M, Thorne N, Wettenhall J. LIMMA: linear models for microarray data. In Bioinformatics and Computational Biology Solutions Using R and Bioconductor. Statistics for Biology and Health. Springer: New York, NY. 2005.

[21] Dusa A, Dusa MA, et al. Package 'venn'. 2021. Available at: https://cran.r-project.org/web/packages/venn/venn.pdf (Accessed: 15 March 2021).

[22] Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, *et al*. STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Research. 2015; 43: D447–D452.

[23] Bader GD, Hogue CW. An automated method for finding molecular complexes in large protein interaction networks. BMC Bioinformatics. 2003; 4: 2.

[24] Chin C, Chen S, Wu H, Ho C, Ko M, Lin C. CytoHubba: identifying hub objects and sub-networks from complex interactome. BMC Systems Biology. 2014; 8: S11.

[25] Yu G, Wang L, Han Y, He Q. ClusterProfiler: an R Package for Comparing Biological Themes among Gene Clusters. OMICS: a Journal of Integrative Biology. 2012; 16: 284–287.

[26] Stretch C, Khan S, Asgarian N, Eisner R, Vaisipour S, Damaraju S, *et al*. Effects of sample size on differential gene expression, rank order and prediction accuracy of a gene signature. PLoS ONE. 2013; 8: e65380.

[27] Hu R, Wang M, Niu W, Wang Y, Liu Y, Liu L, *et al*. SKA3 promotes cell proliferation and migration in cervical cancer by activating the PI3K/Akt signaling pathway. Cancer Cell International. 2018; 18: 183.

[28] Shan L, Zhao M, Lu Y, Ning H, Yang S, Song Y, *et al*. CENPE promotes lung adenocarcinoma proliferation and is directly regulated by FOXM1. Corrigendum in/10.3892/ijo. 2019.4872. International Journal of Oncology. 2019; 55: 257–266.

[29] Wang Z, Yang J, Jiang B, Di J, Gao P, Peng L, *et al*. KIF14 promotes cell proliferation via activation of Akt and is directly targeted by miR-200c in colorectal cancer. International Journal of Oncology. 2018; 53: 1939–1952.

[30] Jing K, Mao Q, Ma P. Decreased expression of TROAP suppresses cellular proliferation, migration and invasion in gastric cancer. Molecular Medicine Reports. 2018; 18: 3020–3026.

[31] Ishitsuka Y, Kawachi Y, Taguchi S, Maruyama H, Fujisawa Y, Furuta J, *et al*. Pituitary tumor-transforming gene 1 enhances proliferation and suppresses early differentiation of keratinocytes. The Journal of Investigative Dermatology. 2012; 132: 1775–1784.

[32] Tsimaratou K, Kletsas D, Kastrinakis NG, Tsantoulis PK, Evangelou K, Sideridou M, *et al*. Evaluation of claspin as a proliferation marker in human cancer and normal tissues. The Journal of Pathology. 2007; 211: 331–339.

[33] Fang H, Niu K, Mo D, Zhu Y, Tan Q, Wei D, *et al*. RecQL4-Aurora B kinase axis is essential for cellular proliferation, cell cycle progression, and mitotic integrity. Oncogenesis. 2018; 7: 68.

[34] Li M, Li A, Zhou S, Lv H, Yang W. SPAG5 upregulation contributes to enhanced c-MYC transcriptional activity via interaction with c-MYC binding protein in triple-negative breast cancer. Journal of Hematology & Oncology. 2019; 12: 14.

[35] Gan Y, Li Y, Li T, Shu G, Yin G. CCNA2 acts as a novel biomarker in regulating the growth and apoptosis of colorectal cancer. Cancer Management and Research. 2018; 10: 5113–5124.

[36] Herring JA, Elison WS, Tessem JS. Function of Nr4a orphan nuclear receptors in proliferation, apoptosis and fuel utilization across tissues. Cells. 2019; 8: 1373.

[37] Tong X, O'Kelly J, Xie D, Mori A, Lemp N, McKenna R, *et*

al. Cyr61 suppresses the growth of non-small-cell lung cancer cells via the β-catenin–c-myc–p53 pathway. Oncogene. 2004; 23: 4847–4855.
[38] Brown JR, Nigh E, Lee RJ, Ye H, Thompson MA, Saudou F, *et al*. Fos family members induce cell cycle entry by activating cyclin D1. Molecular and Cellular Biology. 1998; 18: 5609–5619.
[39] Shaulian E, Karin M. AP-1 in cell proliferation and survival. Oncogene. 2001; 20: 2390–2400.
[40] Yuniati L, Scheijen B, Meer LT, Leeuwen FN. Tumor suppressors BTG1 and BTG2: beyond growth control. Journal of Cellular Physiology. 2019; 234: 5379–5389.

Supplementary material: Supplementary material associated with this article can be found, in the online version, at https://www.imrpress.com/journal/FBL/26/12/10.52586/5038.

Abbreviations: PNUYH, pusan national university yangsan hospital; GO, gene ontology; AEs, acute exacerbations; IPF, idiopathic pulmonary fibrosis; DEGs, differentially expressed genes; PPI, protein–protein interaction; RNA-seq, RNA sequencing; GEO, Gene Expression Omnibus; MCODE, molecular complex detection; STRING, Search Tool for the Retrieval of Interacting Genes; BP, biological process; CC, cellular component; MF, molecular function; BTG1, BTG anti-proliferation factor 1; CCNA2, cyclin A2.

Keywords: IPF; Acute exacerbation; Transplantation; RNA sequencing

Send correspondence to:

Woo Hyun Cho, Division of Allergy, Pulmonary and Critical care medicine, Department of Internal Medicine, Pusan National University Yangsan Hospital, 626-770 Yangsan, Republic of Korea, Research Institute for Convergence of Biomedical Science and Technology, Pusan National University Yangsan Hospital, 626-770 Yangsan, Republic of Korea, E-mail: chowh@pusan.ac.kr

Yun Hak Kim, Department of Anatomy, School of Medicine, Pusan National University, 50612 Yangsan, Republic of Korea, Department of Biomedical Informatics, School of Medicine, Pusan National University, 50612 Yangsan, Republic of Korea, E-mail: yunhak10510@pusan.ac.kr
†These authors contributed equally.