原创精选
Selected article

转化生长因子-β1致肺成纤维细胞转分化过程中差异表达基因的筛选和验证

金路恒[1]，杨培琰[1]，赵阿会[1]，田志鹏[2]，姚武[1]，翟若南[1]，郝长付[1]

1. 郑州大学公共卫生学院尘肺病教研室，河南 郑州 450000
2. 郑州大学第五附属医院福华北街社区服务中心，河南 郑州 450000

摘要：

[背景] 肺成纤维细胞向肌成纤维细胞转分化是矽肺由炎症期进入纤维化期的标志性细胞事件，转化生长因子-β1（TGF-β1）是此环节中最重要的细胞因子。近年来，随着高通量测序和芯片技术的发展，越来越多TGF-β1刺激成纤维细胞转分化的数据被人们所关注。

[目的] 利用生物信息学相关方法探索TGF-β1刺激的人肺成纤维细胞转分化过程中差异共表达基因，筛选与肺成纤维细胞转分化密切相关的转录因子。

[方法] 从基因表达综合数据库GEO检索并下载TGF-β1与肺成纤维细胞相关的高通量数据集GSE17518、GSE97829和GSE119007，对数据集进行差异分析，筛选三组数据集共同上调和共同下调的差异表达基因；对筛选出的差异基因进行基因本体论（GO）和京都基因与基因组百科全书（KEGG）信号通路富集分析；通过HumanTFDB数据库，下载所有人转录因子基因名称和序列，并与上述获得的共表达差异基因进行比对，得到共表达差异转录因子；随后用TGF-β1刺激人胚肺成纤维细胞MRC-5构建转分化模型，对筛选出的差异转录因子进行初步验证。

[结果] 基于这三个数据集，共得到145个上调差异表达基因，88个下调差异表达基因。GO分析结果显示：上述共表达差异基因主要参与细胞外基质、肌动蛋白细胞骨架、内质网腔等组分的构成；并且与肌动蛋白结合、生长因子活性、细胞外基质结构性成分、胶原结合、TGF-β1受体结合等生物学过程相关；此外还与细胞外基质组织、细胞迁移和凋亡，以及丝裂原活化蛋白激酶（MAPK）通路活化、MAPK激活、肽基酪氨酸磷酸化的正调节等有关。通过KEGG信号通路富集分析，筛选出的主要相关信号通路包括TGF-β1信号通路、细胞外基质-受体相互作用、黏着斑和磷脂酰肌醇3激酶-蛋白激酶B（PI3K-Akt）信号通路等。将上述得到的差异基因与人转录因子序列进行比对，得到了7个上调的转录因子和11个下调的转录因子，从中挑选的上调转录因子早期生长应答因子2（EGR2）、SNAIL家族转录抑制因子1（SNAIL1）和下调转录因子胸腺选择相关高迁移转录因子（TOX）、过氧化物酶体增殖物激活受体γ（PPARG）在TGF-β1刺激MRC-5细胞中的qRT-PCR验证结果显示，EGR2和SNAIL1均高表达，而TOX和PPARG均低表达。

[结论] 筛选出18个与肺成纤维细胞转分化相关的转录因子基因，其中在TGF-β1刺激MRC-5细胞中EGR2和SNAIL1表达升高，TOX和PPARG表达降低。

关键词：矽肺；肺成纤维细胞；转化生长因子-β1；差异表达基因；转录因子；生物信息学

Screening and identification of differentially expressed genes in transforming growth factor-β1-induced trans-differentiation of pulmonary fibroblasts

JIN Lu-heng[1], YANG Pei-yan[1], ZHAO A-hui[1], TIAN Zhi-peng[2], YAO Wu[1], ZHAI Ruo-nan[1], HAO Chang-fu[1] (1.Department of Pneumoconiosis, College of Public Health, Zhengzhou University, Zhengzhou, Henan 450000, China; 2.Fuhua North Street Community Service Center, the Fifth Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan 450000, China)

Abstract:

[Background] Pulmonary fibroblast to myofibroblast trans-differentiation is a critical symbol of silicosis progression from inflammation to fibrosis, among which transforming growth factor-beta 1 (TGF-β1) is the most important cytokine. In recent years, with the development of high-throughput sequencing and microarray, more and more TGF-β1-induced fibroblast trans-

DOI 10.13213/j.cnki.jeom.2020.19589

基金项目
国家自然科学基金项目（81472954）

作者简介
金路恒（1994—），女，硕士生；
E-mail：857852873@qq.com

通信作者
郝长付，E-mail：haochangfu@126.com

利益冲突 无申报
收稿日期 2019-08-28
录用日期 2019-11-13

文章编号 2095-9982(2020)03-0211-07
中图分类号 R135
文献标志码 A

▸引用
金路恒，杨培琰，赵阿会，等.转化生长因子-β1致肺成纤维细胞转分化过程中差异表达基因的筛选和验证[J].环境与职业医学，2020，37(3)：211-217.

▸本文链接
www.jeom.org/article/cn/10.13213/j.cnki.jeom.2020.19589

Funding
This study was funded.

Correspondence to
HAO Chang-fu, E-mail: haochangfu@126.com

Competing interests None declared
Received 2019-08-28
Accepted 2019-11-13

▸To cite
JIN Lu-heng, YANG Pei-yan, ZHAO A-hui, et al. Screening and identification of differentially expressed genes in transforming growth factor-β1-induced trans-differentiation of pulmonary fibroblasts[J]. Journal of Environmental and Occupational Medicine, 2020, 37(3): 211-217.

▸Link to this article
www.jeom.org/article/en/10.13213/j.cnki.jeom.2020.19589

differentiation data have been concerned.

[Objective] The current study is designed to explore and screen differentially expressed genes (DEGs) and hub transcription factors (TFs) in trans-differentiation of pulmonary fibroblasts induced by TGF-β1 through bioinformatics strategies.

[Methods] GSE17518, GSE97829, and GSE119007 datasets linking TGF-β1 and pulmonary fibroblasts were downloaded from Gene Expression Omnibus (GEO) database to screen upregulated and downregulated DEGs. Gene Ontology (GO) functional enrichment analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis were performed to the selected DEGs. All obtained DEGs sequences were blast with the TFs sequences downloaded from HumanTFDB database. The selected genes of TFs were validated in TGF-β1 stimulated MRC-5 cells.

[Results] Altogether 145 upregulated DEGs and 88 downregulated DEGs were gained based on the three datasets. The results of GO functional enrichment analysis were mostly enriched in extracellular matrix organization, actin cytoskeleton, endoplasmic reticulum cavity, and other components; they were related to biological processes such as actin binding, growth factor activity, extracellular matrix structural components, collagen binding, and TGF-β1 receptor binding; in addition, they were also related to extracellular matrix tissue, cell migration and apoptosis, mitogen-activated protein kinase (MAPK) pathway activation, MAPK activation, and positive regulation of peptide tyrosine phosphorylation. The results of KEGG pathway enrichment analysis were significantly enriched in TGF-β1 signaling pathway, extracellular matrix-receptor interaction, focal adhesion, and phosphoinositide 3-kinase-protein kinase B (PI3K-Akt) signaling pathway, etc. After comparing with the sequences downloaded from HumanTFDB, 7 upregulated TFs and 11 downregulated TFs were obtained. Early growth response 2 (EGR2), SNAIL family transcriptional repressor 1 (SNAIL1), thymocyte selection-associated high mobility group box protein (TOX), and peroxisome proliferators-activated receptor gamma (PPARG) were validated by qRT-PCR in TGF-β1 stimulated MRC-5 cells, and the results showed that EGR2 and SNAIL1 were upregulated, and TOX and PPARG were downregulated.

[Conclusion] Among the 18 screened genes of TFs related to pulmonary fibrosis trans-differentiation, EGR2 and SNAIL1 are upregulated, and TOX and PPARG are downregulated in TGF-β1 stimulated MRC-5 cells.

Keywords: silicosis; lung fibroblast; transforming growth factor-beta 1; differentially expressed gene; transcription factor; bioinformatics analysis

尘肺病是我国发病率最高且危害最为严重的职业病，尘肺病主要病理表现为肺组织纤维化，迄今尚无可靠的治疗手段[1]。矽肺作为尘肺病中进展最快、死亡率最高的一类疾病，主要是由于长期吸入含有游离二氧化硅的粉尘所致。当粉尘进入肺内，首先被巨噬细胞识别和吞噬，之后巨噬细胞产生大量的致纤维化的细胞因子，如转化生长因子（transforming growth factor beta 1，TGF-β1）、肿瘤坏死因子（tumor necrosis factor-α，TNF-α）等作用于肺成纤维细胞[2]，促使成纤维细胞向肌成纤维细胞转分化[3]。已有研究发现肌成纤维细胞是导致肺组织纤维化的主要终端效应细胞[4]，而作为致纤维化最强的致纤维化因子，TGF-β1 促进肺成纤维细胞向肺肌成纤维细胞转分化这一过程对肺纤维化和矽肺发生发展具有重要意义。

近年来，随着高通量测序数据的发展，为成纤维细胞转分化的研究提供了丰富的测序数据，但因为测序平台、样本选择、作用条件的不同，导致出现很多假阳性结果，难以较为全面地研究成纤维细胞转分化。因此本研究利用生物信息学方法，通过选取来源不同的肺成纤维细胞，对前人学者的高通量测序或芯片数据结果综合分析[5-7]，筛选出共有的应答基因与通路，以期更深入地研究肺成纤维细胞转分化过程，为探索肺纤维化和尘肺的发病机制提供相关参考。

1 材料与方法

1.1 相关基因数据集的选择

从基因表达综合数据库 GEO（http://www.ncbi.nlm.nih.gov/geo）检索和下载 TGF-β1 与肺成纤维细胞相关的高通量数据集 GSE17518、GSE97829 和 GSE119007，其中 GSE17518 对应的是 TGF-β1 刺激肺成纤维细胞 IMR-90 细胞系[5]，GSE97829 对应的是 TGF-β1 刺激肺成纤维细胞 MRC-5 细胞系[6]，GSE119007 对应的是 TGF-β1 刺激原代人肺成纤维细胞[7]。

1.2 筛选差异表达基因

以 $P<0.05$，基因变化倍数的对数值（lg*FC*）<1.2 为标准，通过 R 语言 limma 包对 GSE17518 数据集进行差异分析，R 语言 edgeR 包对 GSE97829 和 GSE119007 进行差异分析，筛选三组数据集差异表达基因，并将这三个数据集得到的差异表达基因取交集，找出共同上调和共同下调的差异表达基因。

1.3 差异表达基因功能注释

将差异表达基因 mRNA 导入 DAVID6.8 数据库（https://david.ncifcrf.gov/conversion.jsp）[8-9]，对筛选出的差异基因进行基因本体论（Gene Ontology，GO）和京都基因与基因组百科全书（Kyoto Encyclopedia of Genes and Genomes，KEGG）信号通路富集分析。

1.4 差异基因中转录因子的筛选

通过 HumanTFDB 数据库（http://bioinfo.life.hust.

edu.cn/HumanTFDB#!/）[10]，下载所有人转录因子基因名称和序列，并与上述获得的共表达差异基因进行比对，得到差异共表达转录因子。

1.5 细胞转分化模型构建

人胚肺成纤维细胞MRC-5购自中国科学院上海细胞库，培养于含10%胎牛血清、100 U·mL^{-1}青霉素、0.1 mg·mL^{-1}链霉素的MEM培养基（中国索莱宝）中。将细胞按照2×10^{5}个·孔$^{-1}$的密度铺入6孔板中，分为实验组和对照组，每组3个平行样本。于37℃、体积分数为5.0%二氧化碳培养箱中稳定培养24 h后，实验组采用浓度为5 ng·mL^{-1}（美国Peprotech）的TGF-β1刺激MRC-5细胞，对照组为未经TGF-β1刺激的MRC-5细胞，48 h后收取RNA样本。

1.6 qRT-PCR法检测基因的表达

用Trizol试剂（日本Takara）提取总RNA，向收集的样品中分别加入200 μL氯仿，混匀静置后12 000×*g*，4℃，离心15 min，分别吸取上清转移至新的离心管中，加入750 μL异丙醇，混匀静置后12 000×*g*，4℃，离心10 min，小心弃去上清，再分别加入1 mL 75%乙醇，7 500×*g*，4℃，离心5 min，弃去上清，得到RNA沉淀。随后使用PrimeScript™ RT reagent Kit with gDNA Eraser试剂盒（日本Takara）对提取的总RNA进行反转录，得到cDNA。最后用TB Green™ Premix Ex Taq II试剂盒（日本Takara）以cDNA为模板进行荧光定量PCR，每组3个重复孔，以*GAPDH*为内参。反应体系：TB Green™ Premix Ex Taq Ⅱ（2×）10 μL，0.4 μL的正向引物，0.4 μL的反向引物，0.4 μL的Rox Reference Dye Ⅱ和6.8 μL的无酶水。反应条件：95℃预变性30 s，95℃变性3 s，60℃退火延伸30 s，共40个循环。所得数值采用$2^{-\Delta\Delta Ct}$计算mRNA的相对表达量。该实验重复3次，引物序列如表1。

表1 qRT-PCR引物序列

Table 1 qRT-PCR primer sequences

基因 Gene	正向引物（5′→3′） Forward	反向引物（5′→3′） Reverse
GAPDH	GCACCGTCAAGGCTGAGAAC	TGGTGAAGACGCCAGTGGA
α-SMA	CTATGAGGGCTATGCCTTGCC	GCTCAGCAGTAGTAACGAAGGA
FN1	AGGAAGCCGAGGTTTTAACTG	AGGACGCTCATAAGTGTCACC
COL1A1	GTGCGATGACGTGATCTGTGA	CGGTGGTTTCTTGGTCGGT
EGR2	TCAACATTGACATGACTGGAGAG	AGTGAAGGTCTGGTTTCTAGGT
SNAIL1	TCGGAAGCCTAACTACAGCGA	AGATGAGCATTGGCAGCGAG
TOX	TATGAGCATGACAGAGCCGAG	GGAAGGAGGAGTAATTGGTGGA
PPARG	ACCAAAGTGCAATCAAAGTGGA	ATGAGGGAGTTGGAAGGCTCT

1.7 统计学分析

采用SPSS 21.0对实验数据进行统计学分析。计量资料经正态性检验符合正态分布时，结果以均数±标准差（$\bar{x}\pm s$）表示。两组独立样本间的均数比较采用*t*检验。多组间均数差异的比较采用单因素方差分析进行比较。检验水准α=0.05。

2 结果

2.1 差异基因的筛选

从GEO数据库中下载GSE17518、GSE97829和GSE119007三个数据集，基于这三个数据集进行差异分析，共筛出233个相关的差异表达基因，其中145个上调差异表达基因，88个下调差异表达基因。

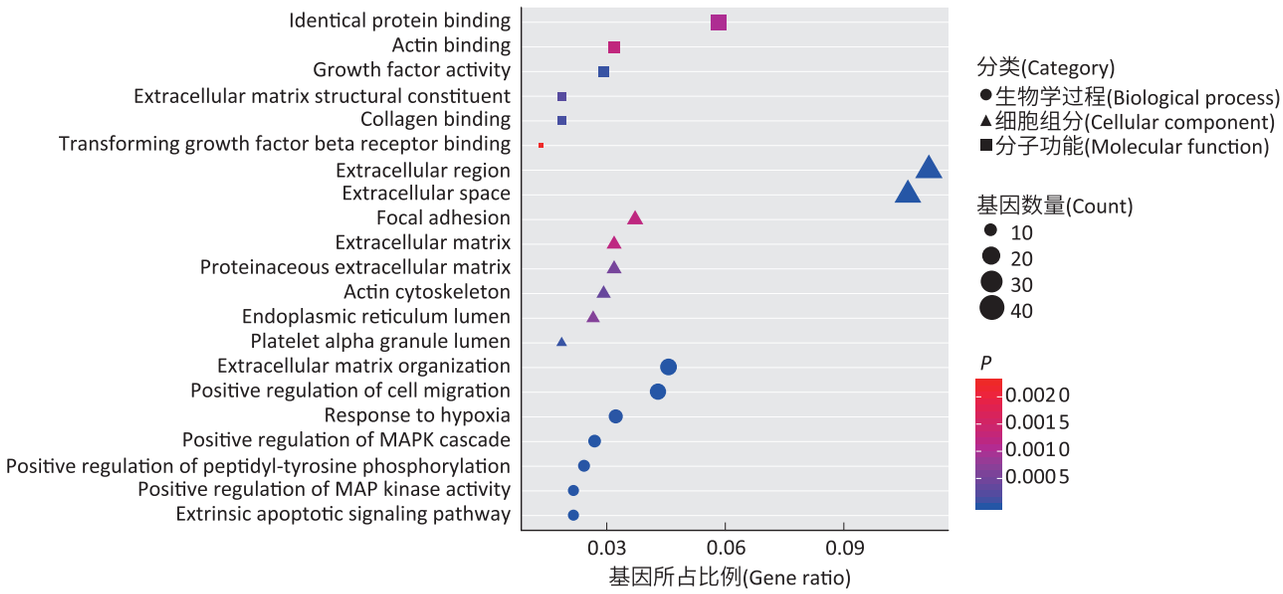
2.2 功能注释

GO分析结果显示，上述差异表达基因主要参与细胞外基质、肌动蛋白细胞骨架、内质网腔等组分的构成；并且与肌动蛋白结合、生长因子活性、细胞外基质结构性成分、胶原结合、TGF-β受体结合等生物学过程相关；此外还与细胞外基质组织、细胞迁移和凋亡，以及丝裂原活化蛋白激酶（mitogen-activated protein kinase，MAPK）通路活化、MAPK激活、肽基酪氨酸磷酸化的正调节等有关（图1）。

通过KEGG信号通路富集分析，筛选出的主要相关信号通路包括TGF-β信号通路、生物合成抗生素、癌症、细胞外基质（extracellular matrixc，ECM）-受体相互作用、黏着斑和磷脂酰肌醇3激酶一蛋白激酶B（phosphoinositide 3 kinase-protein kinase B，PI3K-Akt）信号通路等（图2）。

2.3 转分化相关共表达转录因子筛选

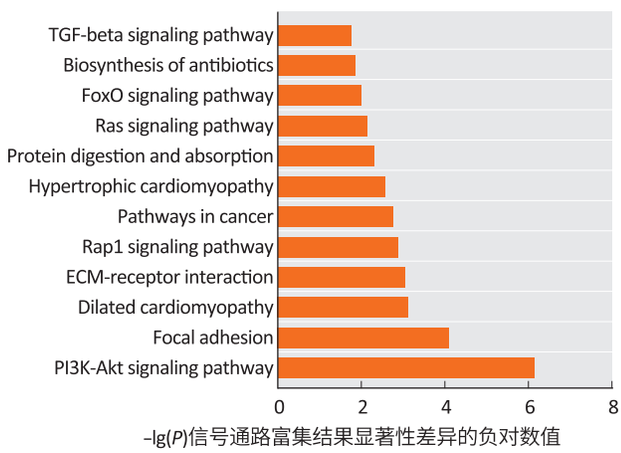
根据GO和KEGG分析结果，我们发现转录调控过程在TGF-β1刺激成纤维细胞转分化过程中发挥了重要作用。因此，我们将上述得到的共表达差异基因与人转录因子序列进行比对，得到了7个上调的转录因子，主要为SMAD7蛋白（SMAD7）、早期生长应答因子2（EGR2）、SNAIL家族转录抑制因子1（SNAIL1）、Krüppel样转录因子7（KLF7）、神经元PAS结构域蛋白2（NPAS2）、基本螺旋-环-螺旋家族成员e40（BHLHE40）和维生素D受体（VDR），11个下调的转录因子主要为T-box转录因子2（TBX2）、SMAD3蛋白（SMAD3）、胸腺选择相关高迁移转录因子（TOX）、NR1H3转录因子、NR2F1转录因子、ZFP36锌指蛋白2（ZFP36L2）、同源异型蛋白SIX1（SIX1）、锌指蛋白395



[注 (Note)] Identical protein binding：相同的蛋白质结合；Actin binding：肌动蛋白结合；Growth factor activity：生长因子活性；Extracellular matrix structural constituent：细胞外基质结构性成分；Collagen binding：胶原结合；Transforming growth factor beta receptor binding：转化生长因子β受体结合；Extracellular region：细胞外区域；Extracellular space：细胞外空间；Focal adhesion：黏着斑；Extracellular matrix：细胞外基质；Proteinaceous extracellular matrix：蛋白质细胞外基质；Actin cytoskeleton：肌动蛋白细胞骨架；Endoplasmic reticulum lumen：内质网腔；Platelet alpha granule lumen：血小板α颗粒腔；Extracellular matrix organization：细胞外基质组织；Positive regulation of cell migration：细胞迁移的正调节；Response to hypoxia：对缺氧的反应；Positive regulation of MAPK cascade：MAPK级联的正调节；Positive regulation of peptidyl-tyrosine phosphorylation：肽基酪氨酸磷酸化的正调节；Positive regulation of MAPK activity：MAPK活性的正调节；Extrinsic apoptotic signaling pathway：外源性细胞凋亡信号通路。

图1 共表达差异基因的GO功能注释

Figure 1 GO functional annotation of differentially expressed genes



[注 (Note)] TGF-beta signaling pathway：TGF-β信号通路；Biosynthesis of antibiotics：抗生素的生物合成；FoxO signaling pathway：FoxO信号通路；Ras signaling pathway：Ras信号通路；Protein digestion and absorption：蛋白质消化和吸收；Hypertrophic cardiomyopathy：肥厚性心肌病；Pathways in cancer：癌症途径；Rap1 signaling pathway：Rap1信号通路；ECM-receptor interaction：ECM-受体相互作用；Dilated cardiomyopathy：扩张型心肌病；Focal adhesion：黏着斑；PI3K-Akt signaling pathway：PI3K-Akt信号通路。

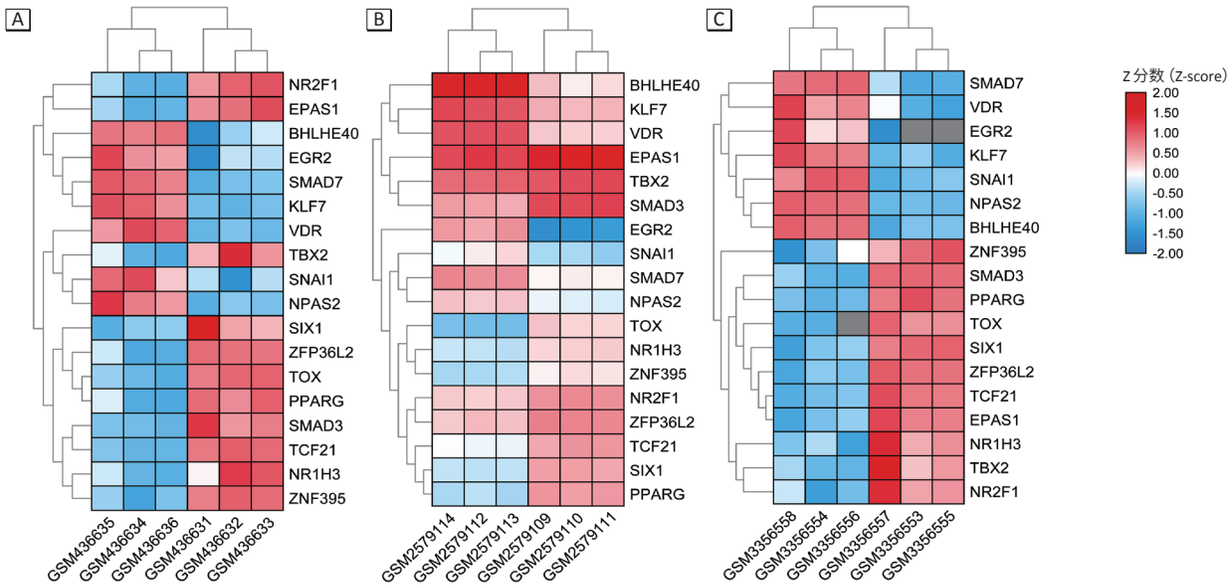
图2 共表达差异基因的KEGG信号通路

Figure 2 KEGG pathway of differentially expressed genes

(ZNF395)、过氧化物酶体增殖物激活受体γ (PPARG)、转录因子21 (TCF21) 和内皮细胞含PAS结构域蛋白-1 (EPAS1)。此外，还得到这些转录因子基因在各数据集中的聚类表达热图，在TGF-β1刺激成纤维细胞转分化后上述转录因子的表达情况一致 (图3)。

2.4 转录因子mRNA验证

用TGF-β1刺激MRC-5细胞，构建肺成纤维细胞转分化的细胞模型，并采用qRT-PCR检测平滑肌肌动蛋白 (alpha-smooth muscle actin alpha，α-SMA)、纤连蛋白1 (fibronectin 1，FN1) 和Ⅰ型胶原α1 (collagen type 1 alpha 1，COL1A1) 的mRNA表达情况以验证肺成纤维细胞转分化模型是否构建成功，qRT-PCR结果显示α-SMA、FN1和COL1A1均高表达 (图4A)。接着从上述筛选出的18个转录因子中，以测序结果中基因变化倍数较明显的上调转录因子EGR2、SNAIL1和下调转录因子TOX、PPARG在MRC-5细胞中进行qRT-PCR验证，发现EGR2和SNAIL1均高表达，而TOX和PPARG均低表达 (图4B)。

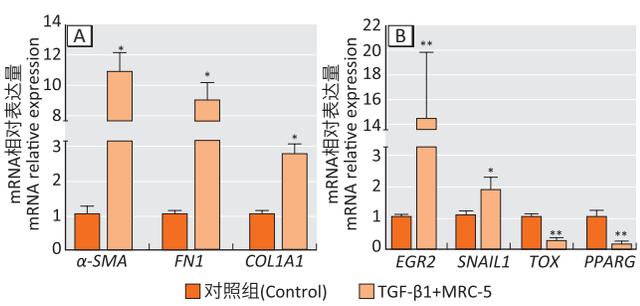


［注］A：GSM436631-GSM436633 为 IMR-90 对照，GSM436634-GSM436636 为 TGF-β1+IMR-90。B：GSM2579109-GSM2579111 为 MRC-5 对照，GSM2579112-GSM2579114 为 TGF-β1+MRC-5。C：GSM3356553、GSM3356555 和 GSM3356557 为肺原代成纤维细胞对照，GSM3356554、GSM3356556 和 GSM3356558 为 TGF-β1+ 肺原代成纤维细胞。SMAD7：SMAD7 蛋白；EGR2：早期生长应答因子 2；SNAIL1：SNAIL 家族转录抑制因子 1；KLF7：Krüppel 样转录因子 7；NPAS2：神经元 PAS 结构域蛋白 2；BHLHE40：基本螺旋 - 环 - 螺旋家族成员 e40；VDR：维生素 D 受体；TBX2：T-box 转录因子 2；SMAD3：SMAD3 蛋白；TOX：胸腺选择相关高迁移转录因子；NR1H3：NR1H3 转录因子；NR2F1：NR2F1 转录因子；ZFP36L2：ZFP36 锌指蛋白 2；SIX1：同源异型蛋白 SIX1；ZNF395：锌指蛋白 395；PPARG：过氧化物酶体增殖物激活受体 γ；TCF21：转录因子 21；EPAS1：内皮细胞含 PAS 结构域蛋白 -1。

［Note］A: GSM436631-GSM436633 is IMR-90 control, and GSM436634-GSM436636 is TGF-β1+IMR-90. B: GSM2579109-GSM2579111 is MRC-5 control, and GSM2579112-GSM2579114 is TGF-β1+MRC-5. C: GSM3356553, GSM3356555, and GSM3356557 are primary pulmonary fibroblast control, and GSM3356554, GSM3356556, and GSM3356558 are TGF-β1+primary pulmonary fibroblast. SMAD7: SMAD7 protein; EGR2: early growth response 2; SNAIL1: SNAIL family transcription repressor 1; KLF7: Krüppel-like transcription factor 7; NPAS2: neuronal PAS domain protein 2; BHLHE40: basic helix-loop-helix family member E40; VDR: vitamin D receptor; TBX2: T-box transcription factor 2; SMAD3: SMAD3 protein; TOX: thymocyte selection-associated high mobility group box protein; NR1H3: NR1H3 transcription factor; NR2F1: NR2F1 transcription factor; ZFP36L2: ZFP36 zinc finger protein 2; SIX1: homeotic protein SIX1; ZNF395: zinc finger protein 395; PPARG: peroxisome proliferators-activated receptor gamma; TCF21: transcription factor 21; EPAS1: endothelial PAS domain-containing protein 1.

图 3　转录因子基因在三个数据集中的表达聚类热图

Figure 3　Heatmap of expression of transcription factors in three datasets



［注］*：$P<0.05$；**：$P<0.001$。*α-SMA*：平滑肌肌动蛋白；*FN1*：纤连蛋白 1；*COL1A1*：Ⅰ型胶原 α1；*EGR2*：早期生长应答因子 2；*SNAIL1*：SNAIL 家族转录抑制因子 1；*TOX*：胸腺选择相关高迁移转录因子；*PPARG*：过氧化物酶体增殖物激活受体 γ。

［Note］*: $P<0.05$; **: $P<0.001$. *α-SMA*: smooth muscle actin alpha; *FN1*: fibronectin 1; *COL1A1*: collagen type I alpha 1; *EGR2*: early growth response 2; *SNAIL1*: SNAIL family transcription repressor 1; *TOX*: thymocyte selection-associated high mobility group box protein; *PPARG:* peroxisome proliferators-activated receptor gamma.

图 4　qRT-PCR 检测各基因 mRNA 的表达情况

Figure 4　mRNA expressions of genes by qRT-PCR

3　讨论

尘肺病是我国目前发病最多且危害最为严重的职业病，2018 年中国共报告各类职业病新病例 23 497 例，其中职业性尘肺病 19 468 例，占 2018 年新发职业病的 82.85%[11]。目前公认的尘肺发病需经历肺泡炎症期、肺部肉芽肿期和纤维化期，而肺成纤维细胞在各种炎症和促纤维化信号的作用下发生转分化，从而转变为导致肺组织纤维化的主要终端效应细胞 —— 肌成纤维细胞[12]。因此从肺成纤维细胞转分化为肌成纤维细胞这一途径入手，对尘肺和纤维化的发生发展具有重要意义。TGF-β1 是重要促纤维化因子，可以激活肺成纤维细胞，使成纤维细胞转分化为肌成纤维细胞[13]，其表达量在二氧化硅导致的肺部损伤修复过程中增加，主要促进肺成纤维细胞过度增殖和迁移[14]，并使其过度释放 ECM、FN1 和 COL1A1，以致正常肺部组织不断地被纤维化的组织取代，严重影响了

尘肺病患者的生存质量。

本研究基于前人研究TGF-β1刺激健康成人肺原代成纤维细胞、IMR-90和MRC-5细胞系的高通量测序或芯片结果，探索这三种肺成纤维细胞转分化过程中均出现上调或下调的差异表达基因，共得到了145个共同上调和88个共同下调的基因。通过生物信息学方法，探索肺成纤维细胞转分化过程中可能涉及的生物学过程和信号通路，发现细胞外基质、肌动蛋白细胞骨架、启动子转录调控和TGF-β1受体结合等为主要的生物学过程。TGF-β1信号通路、ECM-受体相互作用、黏着斑和PI3K-Akt等信号通路在此过程发挥重要作用。

值得注意的是，我们发现转录调控在肺成纤维细胞转分化过程中有很重要的作用。真核生物的转录调控过程非常复杂且具有物种保守性，而转录因子作为基因表达调控的关键分子[15]，主要通过以下几个方面调控基因表达，包括：1. 稳固或阻遏 RNA 聚合酶结合到DNA；2. 催化组蛋白的乙酰化或去乙酰化；3. 招募助活化剂或抑制剂到转录因子（transcription factors，TF）-DNA复合体[16]。虽然转录因子主要以蛋白形式发挥作用，但通过高通量测序结果查找这些转录因子基因的表达情况，可为探索肺成纤维细胞转分化过程中涉及的转录因子所发挥的作用提供参考。因此我们通过将差异共表达基因与人所有转录因子序列进行比对，发现了7个上调和11个下调的转录因子基因。为了进一步验证这些结果，我们构建了TGF-β1刺激MRC-5转分化的细胞模型，并通过qRT-PCR的方法检测到*α-SMA*、*FN1*和*COL1A1*基因均表达增加，证实MRC-5转分化细胞模型构建成功。然后根据高通量测序结果挑选了2个高表达的转录因子基因*EGR2*和*SNAIL1*，2个低表达的转录因子基因*TOX*和*PPARG*进行验证，qRT-PCR结果与测序结果一致。

虽然EGR2蛋白在结构和功能上与早期生长应答因子1（EGR1）相似，主要在周围神经髓鞘形成、脂肪生成和免疫耐受中起重要作用[17]，但Fang等[18]研究发现在博莱霉素诱导的硬皮病小鼠的损伤组织中EGR2水平升高，EGR2过表达可促进成纤维细胞转分化，且TGF-β1可靶向作用于EGR2来促进组织纤维化的发生发展。SNAIL1作为E-粘连蛋白（*E-cad*）基因的上游核心调控因子，主要通过上皮间质转化（epithelial-mesenchymal transition，EMT）过程使上皮细胞转化为肌成纤维细胞[19-20]。杨丽花等[21]发现过氧化氢可诱导MRC-5发生衰老，而衰老的MRC-5细胞分泌的上清里面的炎性因子（白介素-6/白介素-8）可以促进人永生化细胞HaCaT细胞发生恶性转化，从而导致HaCaT细胞发生EMT。PPARG作为一个与脂肪生成相关的转录因子，在TGF-β1刺激肺成纤维细胞转分化过程中低表达，研究表明其能够通过抑制肺成纤维细胞转分化过程从而抑制博来霉素诱导的肺纤维化[22-23]。上述发现与本研究筛选和验证结果相符合，但这些转录因子具体的调控机制以及能否成为肺成纤维细胞转分化过程中重要的节点还有待进一步研究。

本研究从TGF-β1分别刺激三种肺成纤维细胞的高通量转录组数据出发，通过生物信息学方法筛选和挖掘出了18个与肺成纤维细胞转分化密切相关的转录因子基因。本研究的不足之处主要是仅在TGF-β1刺激MRC-5细胞转分化模型中验证相关基因的表达情况，而真核生物转录调控方式十分复杂，本研究未涉及深层次的机制研究，但这些差异表达的转录因子基因可为进一步深入探索肺成纤维细胞转分化转录调控过程提供参考。

参考文献

[1] 谢丽庄，周琅，丁帮梅，等. 2006至2017年某省职业性尘肺病患病病例分析[J]. 中华劳动卫生职业病杂志，2019，37（3）：189-193.

[2] YANG H，LI W，ZHANG Y，et al. Regulatory role of miR-18a to CCN2 by TGF-β1 signaling pathway in pulmonary injury induced by nano-SiO_2[J]. Environ Sci Pollut Res，2018，25（1）：867-876.

[3] 孙林清，王发选，周倩文，等. 转化生长因子-β1对肺成纤维细胞中lncRNA-ATB和纤维化因子mRNA水平的影响[J]. 环境与职业医学，2019，36（3）：261-265.

[4] LI J，YAO W，ZHANG L，et al. Genome-wide DNA methylation analysis in lung fibroblasts co-cultured with silica-exposed alveolar macrophages[J]. Respir Res，2017，18（1）：91.

[5] HECKER L，VITTAL R，JONES T，et al. NADPH oxidase-4 mediates myofibroblast activation and fibrogenic responses to lung injury[J]. Nat Med，2009，15（9）：1077-1081.

[6] SAVARY G，DEWAELES E，DIAZZI S，et al. The long noncoding RNA DNM3OS is a reservoir of FibromiRs with major functions in lung fibroblast response to TGF-β and pulmonary fibrosis[J]. Am J Respir Crit Care Med，2019，200（2）：184-198.

[7] MUKHERJEE S，SHENG W，MICHKOV A，et al. Prostaglandin E2 inhibits profibrotic function of human pulmonary fibroblasts by disrupting Ca^{2+} signaling [J]. Am J Physiol Lung Cell Mol Physiol，2019，316 (5)：L810-L821.

[8] HUANG D W，SHERMAN B T，LEMPICKI R A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources [J]. Nat Protoc，2009，4 (1)：44-57.

[9] HUANG D W，SHERMAN B T，LEMPICKI R A. Bioinformatics enrichment tools：paths toward the comprehensive functional analysis of large gene lists [J]. Nucleic Acids Res，2009，37 (1)：1-13.

[10] HU H，MIAO Y R，JIA L H，et al. AnimalTFDB 3.0：a comprehensive resource for annotation and prediction of animal transcription factors [J]. Nucleic Acids Res，2019，47 (D1)：D33-D38.

[11] 中华人民共和国卫生健康委. 2018年我国卫生健康事业发展统计公报 [EB/OL]. [2019-08-21]. http://www.nhc.gov.cn/guihuaxxs/s10748/201905/9b8d52727cf346049de8acce25ffcbd0.shtml.

[12] VINCENT M，CHEMARIN C，CAVALIN C，et al. From the definition of silicosis at the 1930 Johannesburg conference to the blurred boundaries between pneumoconioses，sarcoidosis，and pulmonary alveolar proteinosis (PAP) [J]. Am J Ind Med，2015，58 (S1)：31-38.

[13] ZHENG X，QI C，ZHANG S，et al. TGF-β1 induces Fstl1 via the Smad3-c-Jun pathway in lung fibroblasts [J]. Am J Physiol Lung Cell Mol Physiol，2017，313 (2)：L240-L251.

[14] 王凯，刘志宏，王和静，等. 氧化铍致大鼠肺纤维化及相关指标的动态变化 [J]. 环境与职业医学，2016，33 (7)：644-649.

[15] DALY M E. Transcription factor defects causing platelet disorders [J]. Blood Rev，2017，31 (1)：1-10.

[16] DAI X G，LI T，HUANG W B，et al. Upregulation of mitochondrial transcription factor a promotes the repairment of renal tubular epithelial cells in sepsis by inhibiting reactive oxygen species-mediated toll-like receptor 4/p38MAPK signaling [J]. Pathobiology，2019，86 (5-6)：263-273.

[17] DOLDER P C，GRÜNBLATT E，MÜLLER F，et al. A single dose of LSD does not alter gene expression of the serotonin 2A receptor gene (HTR2A) or early growth response genes (EGR1-3) in healthy subjects [J]. Front Pharmacol，2017，8：423.

[18] FANG F，OOKA K，BHATTACHARYYA S，et al. The early growth response gene Egr2 (Alias Krox20) is a novel transcriptional target of transforming growth factor-β that is up-regulated in systemic sclerosis and mediates profibrotic responses [J]. Am J Pathol，2011，178 (5)：2077-2090.

[19] LIANG H，LIU S，CHEN Y，et al. miR-26a suppresses EMT by disrupting the Lin28B/let-7d axis：potential cross-talks among miRNAs in IPF [J]. J Mol Med (Berl)，2016，94 (6)：655-665.

[20] SWAIN S D，GRIFKA-WALK H N，GRIPENTROG J，et al. Slug and Snail have differential effects in directing colonic epithelial wound healing and partially mediate the restitutive effects of butyrate [J]. Am J Physiol Gastrointest Liver Physiol，2019，317 (4)：G531-G544.

[21] 杨丽花. 丹皮酚对MRC-5细胞的抗衰老和抑制HaCaT细胞的EMT作用 [D]. 广州：广州中医药大学，2017.

[22] EL AGHA E，MOISEENKO A，KHEIROLLAHI V，et al. Two-way conversion between lipogenic and myogenic fibroblastic phenotypes marks the progression and resolution of lung fibrosis [J]. Cell Stem Cell，2017，20 (4)：261-273.e3.

[23] MILAM J E，KESHAMOUNI V G，PHAN S H，et al. PPAR-γ agonists inhibit profibrotic phenotypes in human lung fibroblasts and bleomycin-induced pulmonary fibrosis [J]. Am J Physiol Lung Cell Mol Physiol，2008，294 (5)：L891-L901.

（**英文编辑**：汪源；**编辑**：王晓宇；**校对**：陈姣）