Temporal Gene Expression of Mesenchymal Cells in the Pediatric Lung

Quinlen F. Marshall,[1,2] Soumyaroop Bhattacharya, M.S.,[3,4] Gautam Bandyopadhyay, Ph.D.,[3]
Ravi Misra, Ph.D.,[3] Thomas Mariani, Ph.D.,[3,4] Gloria Pryhuber, M.D.[3]

[1]*Saint John's University, Collegeville, Minnesota;* [2]*Strong Children's Research Center, University of Rochester Medical Center, Rochester, New York;* [3]*Division of Neonatology, Department of Pediatrics, University of Rochester Medical Center, Rochester, New York;* [4]*Program in Pediatric Molecular and Personalized Medicine, Department of Pediatrics, University of Rochester Medical Center, Rochester, New York.*

INTRODUCTION: The newborn lung undergoes vast biochemical and physiological changes during adaptation from the intrauterine to the extrauterine environment. Lung morphogenesis continues from birth into early childhood, mediated by dynamic gene expression and a diversity of pulmonary cell types (Whitsett, JA. et al. *Physiol. Rev*, 2019). Murine models demonstrate that pulmonary mesenchymal cells exhibit remarkable heterogeneity in function and morphology during development, however, confirmation of their role is lacking in human neonates and early childhood (Guo, M. et al. *Nat. Comm,* 2019). In addition, many current human genomic studies of lung maturation suffer from limited sample size, limiting their applicability to longitudinal pediatric lung development. Temporal analysis of gene expression aims to bridge this gap, and the most common analytical approach utilizes Short Time-series Expression Miner (STEM) (Ernst, J. & Bar-Joseph, Z. *BMC Bioinformatics*, 2006). STEM utilizes unique methods to cluster, compare, and visualize short time-series gene expression data.

METHODS: Dissociation of lung cells, sorting into enriched populations, and RNA isolation was performed at the Human Tissue Core of the Molecular Atlas of Lung Development Program (Bandyopadhyay, G. et al. *Am. J. Physiol. Lung Cell Mol. Physiol,* 2018). RNA sequencing (RNAseq) was performed at the University of Rochester Genomics Research Center using the Ilumina NovaSeq6000, and reads were aligned using the Splice Transcript Alignment to a Reference algorithm (STaR). Reads were further normalized using counts per million (CPM) and variance-mean dependence calculated with DESeq as implemented in Bioconductor. Genes not detected in at least 3 time points or exhibiting a minimum fold change of at least 3 across the time series were excluded from further analysis. Time-series analysis was performed with STEM, and profiles were assigned significance by Fisher's exact test ($p<0.05$). Genes selected from profiles of interest were functionally enriched using ToppGene Functional Gene Enricher (Chen, J. et al. *BMC Bioinformatics*, 2007).

RESULTS: RNAseq was performed using RNA obtained from pulmonary mesenchymal cells, (n=24, (<1 d/o - 8 y/o, 17 m, 7 f) generating 24.3±5.5 million reads at depth of 10 million reads (48.3±4.6% of genome mapped). CPM normalized expression values for repeat donor time points were averaged and then separated into a younger (n=9, <1 d/o - 1 y/o) and older (n=8, 1 y/o - 8 y/o) group. A total of 17,843 genes passed filtering criteria in the younger group and 17,840 passed in the older group. Using STEM, 16 and 20 profiles were found to be significant in the younger and older group, respectively. 7 profiles in the younger group and 8 profiles in the older group were selected for further functional analysis based on significance and directionality of gene expression changes.

Multiple profiles in both groups demonstrated matrix fibroblast associated gene expression increasing in both groups, peaking at 2 years. Next, proliferative fibroblast and cell division associated gene expression decreased from birth to 1 year in the younger group. Detection of multiple mesenchymal-like profiles validates the purity of cells enriched. Additionally, gene expression associated with immune-like pathways increased in both groups. Finally, cell signatures in the older group associated with the Wnt pathway decreased from 1 year until 2 years and then increased from 4 years to 8 years.

CONCLUSIONS: In summary, analysis of dynamic gene expression in isolated cells across a time series demonstrates the unique heterogeneity of pulmonary mesenchymal cells throughout adolescence. In addition, increased gene expression associated with immune signatures during pediatric lung development was noted. Further validation and exploration using this technique may advance understanding of the diversity of pulmonary cell types and pathophysiology of pediatric lung disease.