A Bulk RNA-seq Metric of Transcriptional Plasticity Is Associated with Clinical Outcomes in Pneumonia Patients

Samuel Dadd Hamilton, Vadim Backman, & Deborah Rachelle Winter

Abstract

All cells must respond to their environment to survive, yet a lack of methods to quantify transcriptional plasticity impedes our ability to measure this essential process. In the case of macrophages, a highly adaptable immune cell type, this knowledge gap inhibits our ability to understand how miscalibrated responsiveness contributes to disease. Here, we drew on our previous work showing how *Chromatin Packing Complexity* facilitates transcriptional plasticity through increased *Transcriptional Divergence* and *Intercellular Heterogeneity*, to develop a proxy-metric for macrophage plasticity in bulk RNA-seq data. We found this metric was significantly associated with key clinical metrics in intubated pneumonia patients. Our results suggest this metric could have value in future studies on patient response and macrophage plasticity.

Background

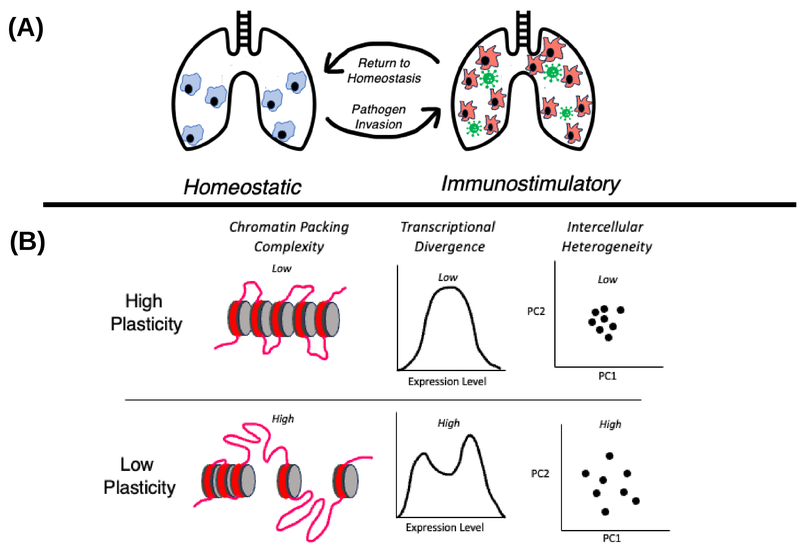
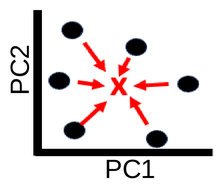


Figure 1: **(A)** Diagram illustrating how macrophages adapt transcription in response to foreign invaders in the lung. **(B)** Diagram describing previous work on the relationship between plasticity, chromatin packing heterogeneity, transcriptional divergence, and intercellular heterogeneity

Materials & Methods

(A) *Intercellular Heterogeneity Calculation*

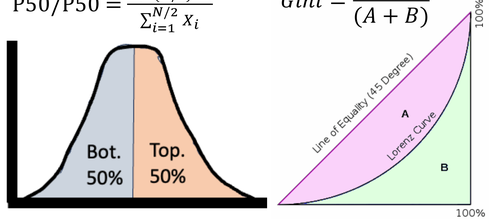
$$\frac{1}{N}\sum_{i=1}^{N} d(x_i, C)$$



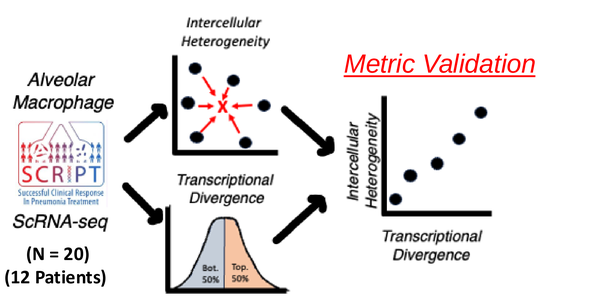
(B) *Transcriptional Divergence Calculation*

$$\text{P50/P50} = \frac{\sum_{i=(N/2)}^{N} X_i}{\sum_{i=1}^{N/2} X_i}$$

$$Gini = \frac{A}{(A+B)}$$



(C)



(D)

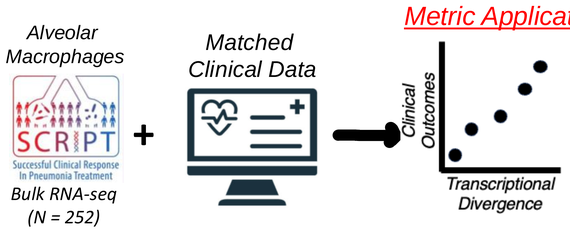


Figure 2: **(A)** Intercellular Heterogeneity was calculated by averaging the distance of all cells to a computed centroid in PCA space. **(B)** Two methods for estimating *Transcriptional Divergence* were employed. Both methods first order genes by their expression level. P50/P50 divides the total expression of the top 50% of genes over the bottom 50%. Gini divides the CDF of gene expression for a sample by the CDF of a hypothetical sample where all reads are distributed evenly. **(C)** Validation Approach: scRNA-seq samples generated from alveolar macrophages from intubated patients enrolled in the Successful Clinical Response in Pneumonia Therapy Project (SCRIPT) were used to calculate *Intercellular Heterogeneity* and *Transcriptional Divergence*. The relationship between these metrics were used to evaluate the transcriptional divergence metric. **(D)** Metric Application: The *Transcriptional Divergence* metric was applied to alveolar macrophage bulk RNA-seq from SCRIPT and correlated with matched clinical data.

Novel Methodology Can Address Common Technical Biases When Calculating Transcriptional Divergence

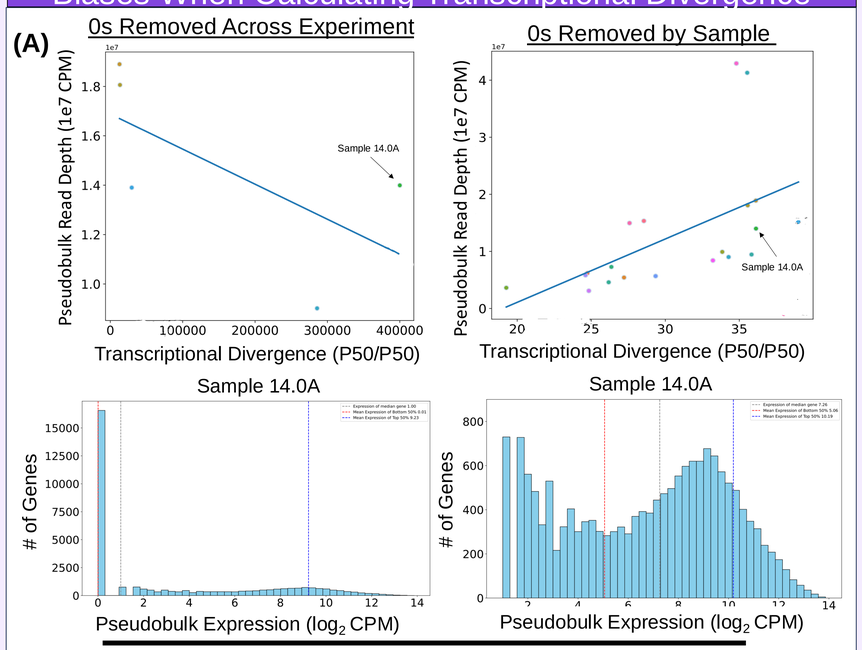


Figure 3: The effects of technical factors such as read depth on transcriptional divergence calculation and how they can be addressed. For scatterplots, dots represent samples and blue lines show the line of best fit. **(A)** Removing genes with 0 detected counts from each sample individually (left) rather than removing those that were not expressed across all samples (left) improves the applicable range of the P50/P50 metric. Histograms show the distribution of gene expression for example sample 14.0A for each 0 removal method. Dashed blue, black, and red lines show the $P50_{bot}$, median, and $P50_{top}$ values respectively **(B)** Standardizing read depth by down-sampling reads can address the effect of read depth on transcriptional divergence. **(C)** Metrics sensitive to low count reads like the Gini Coefficient are more affected by read sparsity than coarse metrics such as P50/P50.

Transcriptional Divergence Is Associated With Intercellular Heterogeneity

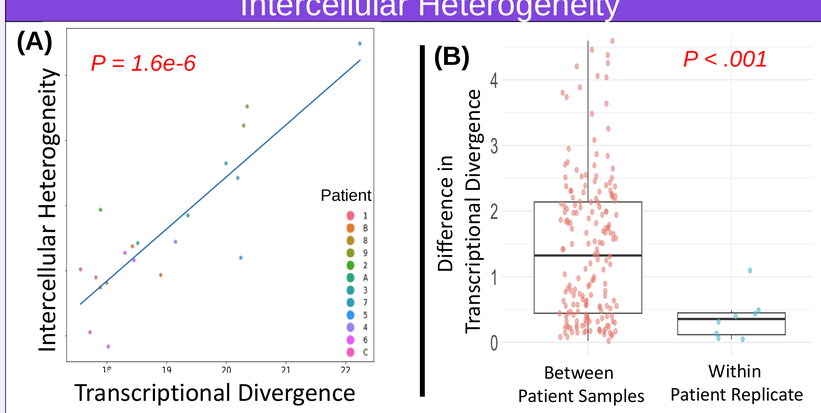


Figure 4: **(A)** The relationship between transcriptional divergence and intercellular heterogeneity, with 0s removed, read depth normalized, using the P50/P50 metric (See Fig. 2). Each dot represents a sample. Dot colors represent the patient from which samples were taken. The blue line represents the line of best fit to the data. **(B)** The robustness of transcriptional divergence calculation across technical replicates. Each point represents the difference in calculated value between two samples. Color indicates the type of relationship, where red are non-replicate differences, and blue are within replicate differences.

Transcriptional Divergence Is Associated With Key Clinical Outcomes In Pneumonia Patients

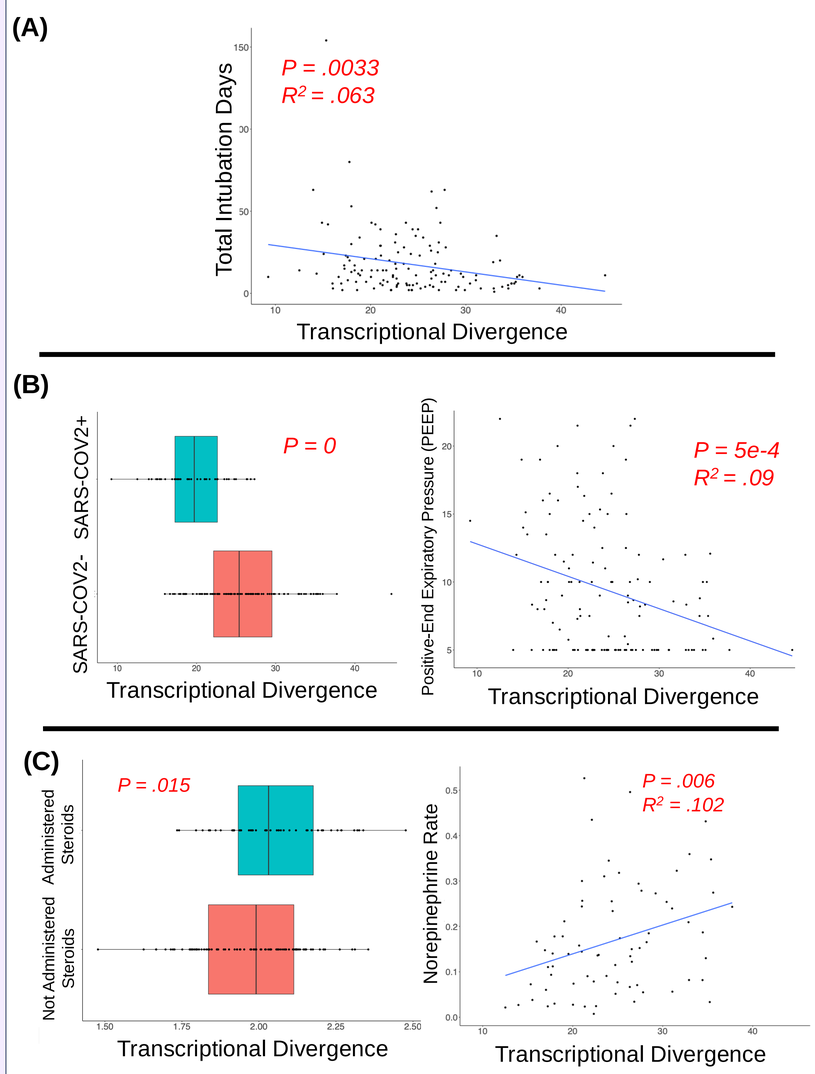


Figure 5: The relationships between transcriptional divergence calculated from bulk RNA-seq and key clinical metrics in intubated pneumonia patients. Each point represents a patient. For continuous response variables, the blue lines represent the line of best fit. For binary response variables, box and whisker plots denote the positions of the minimum, 1st quartile, median, 3rd quartile, and maximum values. P-values for continuous and binary response variables were calculated by fitting a linear and logistic regression functions, respectively.**(A)** Patients with lower transcriptional divergence were intubated longer. **(B)** Patients with low transcriptional divergence were more likely to have been hospitalized with SARS-COV2 and were given greater ventilator support. The PEEP value is the average value on the day the bulk RNA-seq sample was collected. **(C)** Patients with high transcriptional divergence were more likely to be given anti-inflammatory treatment, and with greater intensity. The norepinephrine rate is the average value on the day the bulk RNA-seq sample was collected. Patients that were not given norepinephrine were not included. **(D)** Patients with high transcriptional divergence expressed more common inflammatory markers.

Conclusions

- Transcriptional divergence can be calculated robustly and reliably with methodology that minimizes the effect of read depth
- Transcriptional divergence was reliably associated with intercellular heterogeneity in alveolar macrophages from pneumonia patients
- The relationship between macrophage transcriptional divergence and clinical metrics suggest transcriptional divergence may be associated with responsiveness in pneumonia patients.

References & Acknowledgments

This research was supported by funds from the Physical Genomics and Engineering Training Program Grant (T32GM142604)

Grant, R. A., Morales-Nebreda, L., Markov, N. S., Swaminathan, S., Querrey, M., Guzman, E. R., Abbott, D. A., Donnelly, H. K., Donayre, A., Goldberg, I. A., Klug, Z. M., Borkowski, N., Lu, Z., Kihshen, H., Politanska, Y., Sichizya, L., Kang, M., Shilatifard, A., Qi, C., … Wunderink, R. G. (2021). Circuits between infected macrophages and T cells in SARS-CoV-2 pneumonia. *Nature*, *590*(7847), 635–641.

Reyfman, P. A., Walter, J. M., Joshi, N., Anekalla, K. R., McQuattie-Pimentel, A. C., Chiu, S., Fernandez, R., Akbarpour, M., Chen, C.-I., Ren, Z., Verma, R., Abdala-Valencia, H., Nam, K., Chi, M., Han, S., Gonzalez-Gonzalez, F. J., Soberanes, S., Watanabe, S., Williams, K. J. N., … Misharin, A. V. (2019). Single-Cell Transcriptomic Analysis of Human Lung Provides Insights into the Pathobiology of Pulmonary Fibrosis. *American Journal of Respiratory and Critical Care Medicine*, *199*(12), 1517–1536.