#202 Integrating single cell and spatial gene expression profiling of mouse organogenesis to identify and localize unknown cell types

Chengxiang Qiu[1], Eva Nichols[1], Stephanie Zimmerman[2], Alexa Lasley[2], Margaret Hoang[2], Joe Beechem[2], Jay Shendure[1]

1. Department of Genome Sciences, University of Washington, Seattle, WA, USA, 2. NanoString® Technologies, Seattle, WA, USA

NanoString Technologies
530 Fairview Avenue North, Seattle, WA 98109

Abstract

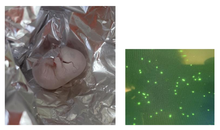
Mammalian organogenesis is a remarkable process, whereby cells within the post-gastrulation embryo continue to rapidly proliferate while giving rise to the diverse cell types of each organ system, specified by molecular programs that are precisely regulated in time and space. Single cell RNA-sequencing of whole embryos during mouse embryogenesis and organogenesis is yielding unprecedented detailed views of mammalian development, for example revealing hundreds of unique cell types defined by gene expression. Although many methods exist for identification of cell types defined by scRNA-seq, annotating cells remains a manual and challenging process. In this work, we sought to leverage spatial gene expression data of mouse organogenesis to validate annotations and localize uncertain cell populations to specific tissues or regions.

We used high-resolution scRNA-seq data from mouse development -- specifically, single nucleus transcriptional profiling of millions of cells done by 3-level combinatorial indexing. This “whole organism profiling” was performed on staged embryos in 2 to 6 hour increments from gastrulation to birth. The resulting single cell profiles were processed by conventional methods and, although an initial round of manual annotation based on marker genes and earlier generations of atlases was fruitful, many ambiguities remained. To address these in part, we integrated matched timepoints with spatial whole transcriptome profiles of specific anatomical structures of four timepoints of mid-gestation mouse development generated using the GeoMx® Digital Spatial Profiler (DSP). We used a cell type deconvolution algorithm to estimate the abundance of each cell type in each region profiled by DSP and validated that known cell types such as tissue-specific epithelial cell subtypes localize to the correct anatomical structures with high accuracy. We then used this method to further map the cell trajectories derived from the lateral plate mesoderm, populations which have limited research and are therefore challenging to annotate.

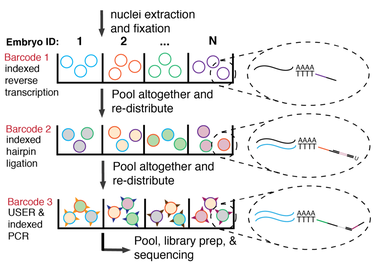
Next, we applied this method to understand how dysregulated cell lineage contributes to organ malformation in a developmental mutant. Absence of embryonic macrophages due to colony stimulating factor 1 receptor (CSF1R) deficiency causes bone and brain deformities as well as perinatal lethality in mouse and humans, suggesting important functions in organ formation. We performed massively scalable RNA single-cell transcriptomics (via single-cell combinatorial indexing RNA sequencing) and GeoMx DSP on E18.5 embryonic tissue sections of wildtype and CSF1R-deficient mutant littermates. We find differential cell type abundance in both the scRNAseq and between matched spatial regions in wild type and mutant in a wide variety of tissues, suggesting that organs beyond bone and brain are impacted by embryonic macrophage loss. In conclusion, this work provides a framework for integrating spatial data with scRNAseq in an automated pipeline to add spatial annotations to unknown cell types in normal and pathological samples.

Methods: single cell RNAseq and spatial whole transcriptome profiling of mouse organogenesis

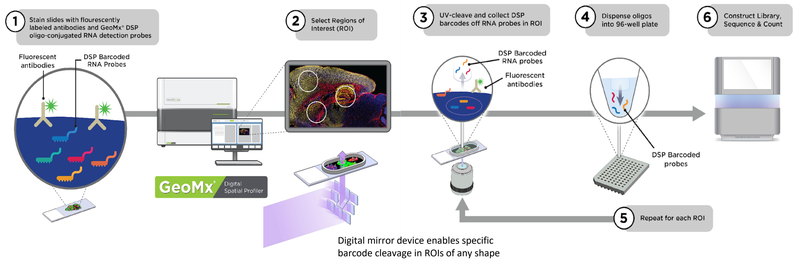
Massively scalable scRNAseq (sci-RNAseq[1,2]) enables high resolution profiling of mouse organogenesis



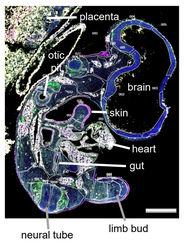
Flash-frozen embryos are pulverized for nuclei isolation and scRNAseq libraries are prepared using a 3-level combinatorial indexing approach. Embryos were profiled at 2-6 hour increments across development and 26 major cell type trajectories identified.



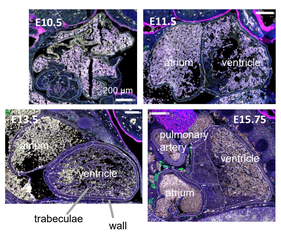
GeoMx Digital Spatial Profiling with the Mouse Whole Transcriptome Atlas[3] enables quantification of 20,175 protein-coding genes in precisely defined regions of interest



Comprehensive profiling at E10.5 with ROIs drawn to capture precise anatomical structures

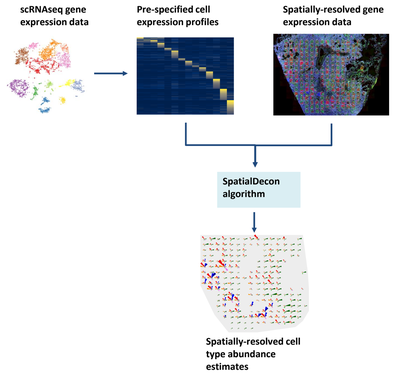


Example regions of interest in heart at all four timepoints



We used a publicly available spatial mouse development dataset from the NanoString Spatial Organ Atlas with four matched timepoints to the wild type sciRNAseq development data (E10.5, E11.5, E13.5, E15.75). These data include spatial whole transcriptome profiling of 15 developing mouse organs, with regions of interest precisely drawn to capture specific histological structures.

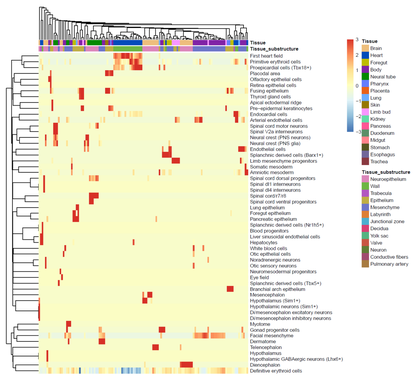
The SpatialDecon[4] algorithm estimates abundances of mixed cell populations in spatially resolved gene expression data using cell profiles derived from scRNAseq and can localize unannotated cell clusters



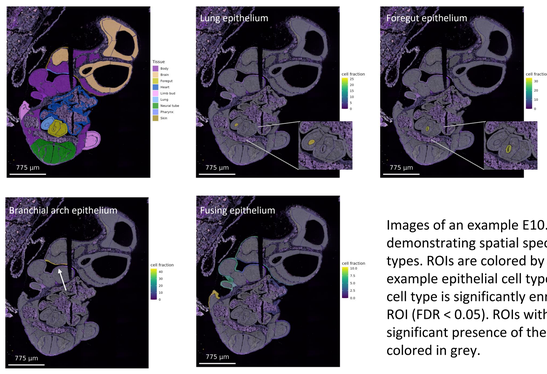
The SpatialDecon algorithm uses log-normal regression to estimate cell abundances in spatial data based on a reference gene expression profile for each cell type, and provides confidence intervals and pvalues for cell abundance estimates. Data from single cell RNAseq can be used to construct reference profiles, with cell types defined by any method of classifying cells into subgroups.

Construction of cell profiles and cell type deconvolution were implemented using the SpatialDecon R package (https://bioconductor.org/packages/release/bioc/html/SpatialDecon.html) and the results spatially visualized using the SpatialOmicsOverlay R package (https://github.com/Nanostring-Biostats/SpatialOmicsOverlay).

Major cell clusters in wild type E10.5 embryo show expected spatial patterns



Major cell clusters in the developmental scRNAseq dataset were mapped to regions of interest in age-matched E10.5 embryos. Heatmap shows enrichment of each scRNAseq cell type in each GeoMx ROI, annotated by tissue and substructure in the GeoMx data. Only cell types significantly enriched in at least one ROI are shown.



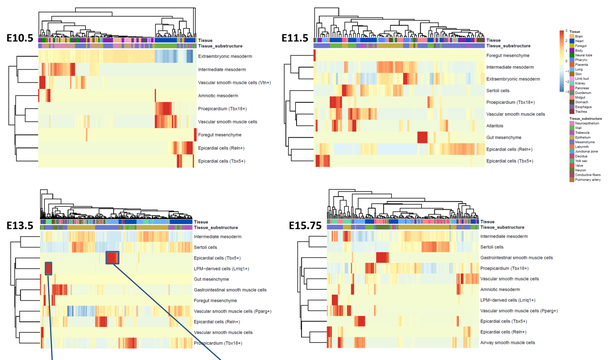
Associated with first branchial arch

Associated with skin

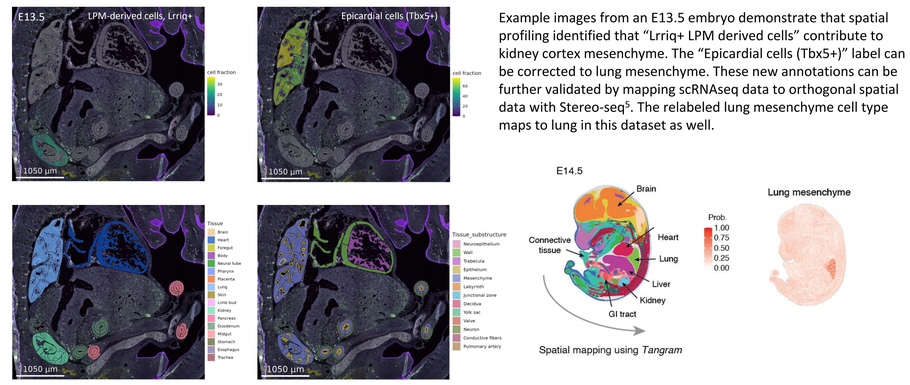
Images of an example E10.5 embryo demonstrating spatial specificity of cell types. ROIs are colored by abundance of example epithelial cell types when that cell type is significantly enriched in the ROI (FDR < 0.05). ROIs with no significant presence of the cell type are colored in grey.

Mapping lateral plate mesoderm trajectories across mid-gestation development

Some previously unannotated cell types can be identified and relabeled based on spatial location



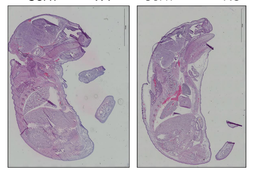
Cells derived from the lateral plate mesoderm contribute to many organs in the developing embryo, but lateral plate mesoderm lineages are complex and difficult to annotate. To validate and correct annotations, we mapped sub-trajectories of lateral plate mesoderm to spatial locations at matched timepoints in mid-gestation development. Heatmaps shows enrichment of each cell type at each timepoint in each GeoMx ROI, annotated by tissue and substructure in the GeoMx data. Only cell types significantly enriched in at least one ROI are shown.



Example images from an E13.5 embryo demonstrate that spatial profiling identified that “Lrriq+ LPM derived cells” contribute to kidney cortex mesenchyme. The “Epicardial cells (Tbx5+)” label can be corrected to lung mesenchyme. These new annotations can be further validated by mapping scRNAseq data to orthogonal spatial data with Stereo-seq[5]. The relabeled lung mesenchyme cell type maps to lung in this dataset as well.

Application to a developmental mutant: sciRNAseq and spatial profiling of *Csfr1* deficient embryos

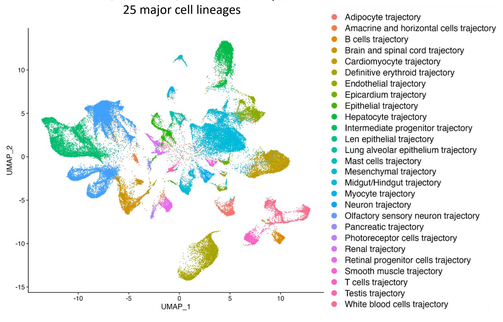
Mouse lacking embryonic macrophages exhibit profound developmental defects in multiple tissues



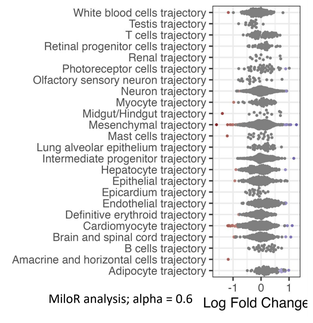
Embryonic macrophages arise from yolk-sac island progenitor cells which migrate to developing organ buds, expanding from E9.5 of mouse organogenesis to become functionally distinct long-lived tissue-resident macrophages. Loss of embryonic macrophages due to *Csfr1* deficiency causes bone and brain deformities as well as perinatal lethality in mouse and humans[6].

Littermate WT and *Csfr1* KO E18.5 embryos were profiled by sciRNAseq and cells were mapped to 25 major trajectories

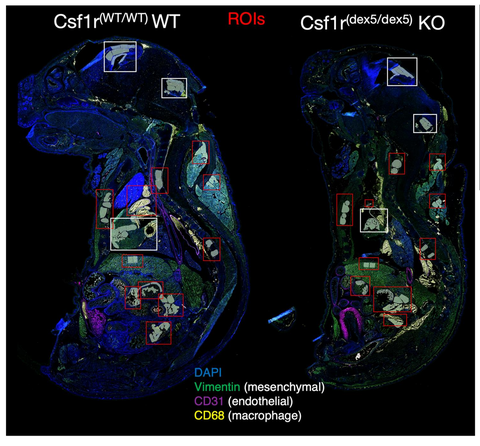
59,095 high-quality cells in one experiment from 3♂ WT and 3♂ KO whole embryos 25 major cell lineages

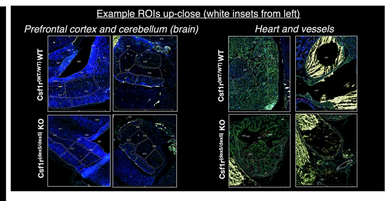


scRNAseq data suggests modest changes in cell abundance

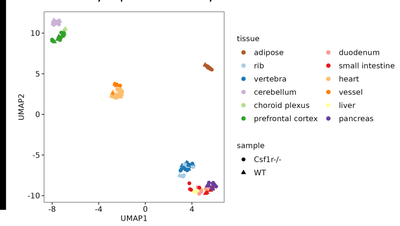


GeoMx Digital Spatial Profiling of 12 major tissues and substructures within those tissues in matched fixed frozen sagittal sections of littermate E18.5 WT and *Csfr1* deficient embryos



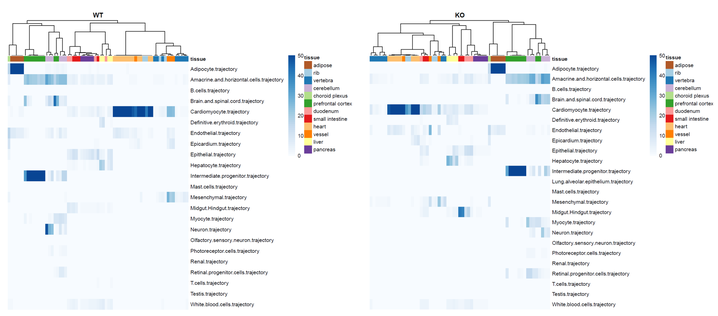


UMAP clearly separates ROIs by tissue

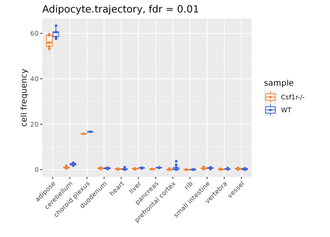


Mapping major cell types spatially and identifying tissue-specific differential cell abundance in *Csfr1* mutants

Abundance of each major trajectory in each region of interest in wild type and mutant embryos using matched sciRNAseq and GeoMx DSP data



Spatial patterning of major trajectories is overall similar between wild type and mutant, with minor differences in abundance in certain populations. For example, the adipocyte trajectory is less abundant in multiple tissues, including adipose and brain. Future work will investigate the spatial distribution and abundance of the sub-lineages of each cell type in wild type and mutant.



Conclusions

- High resolution sciRNAseq identifies novel cell types during dynamic processes such as embryonic development
- Spatial whole transcriptome profiling with GeoMx WTA can localize and annotate novel cell types identified by scRNAseq
- In developmental mutants, single cell RNAseq plus spatial gene expression data can link observed pathology to changes in cell abundance, cell localization, or cell state

References

1. Cao, J. et al. *Nature* 566: 496–502 (2019).
2. Martin, B. K. et al. *Nat Protoc* 18: 188–207 (2023).
3. Zimmerman, S.M. et al. *Genome Res* 32(10): 1892-1905 (2022)
4. Danaher P. et al., *Nat Comm* 13(1): 385 (2022)
5. Chen, A. et al. *Cell* 185(10): 1777-1792 (2022)
6. Hume, D. A. et al. *J Leukoc Biol* 107: 205–219 (2020).

Scan here to download or learn more