Cells2Spine: Mapping the Cellular Landscape of Spinal Cord Injury with Spatial Transcriptomics and Single Cell RNA-Seq

Quique Vidal-Beneyto [1, 2, 3, 4], Loris Mannino [4], Guillem Paniagua-Soriano [4], Victoria Moreno-Manzano [4], Sonia Tarazona [3], Ana Conesa [2]

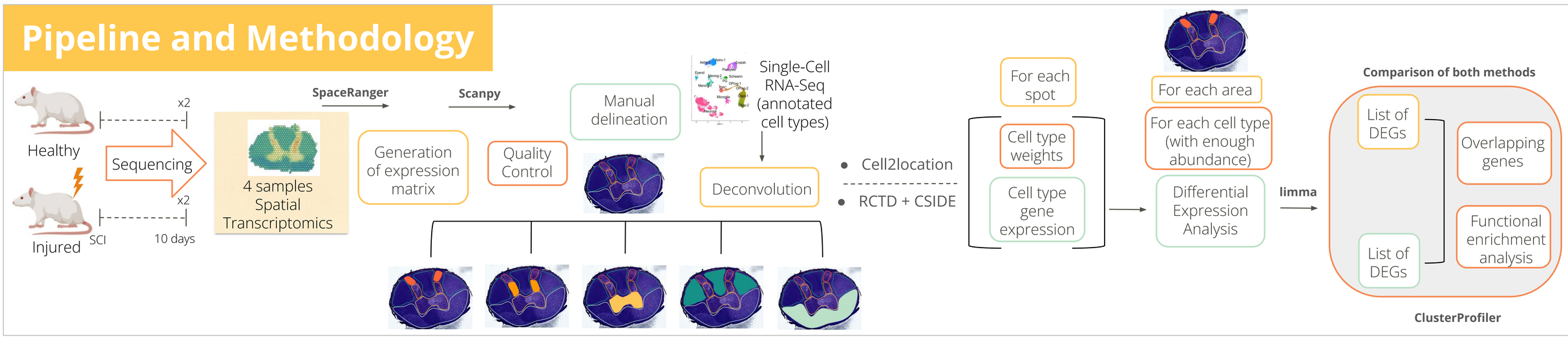
1. PhD program in Biotechnology, Universitat Politècnica de València (UPV), València, Spain
2. Genomics of Gene Expression Lab, Institute for Integrative Systems Biology, Consejo Superior de Investigaciones Científicas (I2SysBio-CISC-UV), Paterna, Spain
3. BiostatOmics Group, Departamento de Estadística e Investigación Operativa, Universitat Politècnica de València (UPV), València, Spain
4. Neuronal and Tissue Regeneration Lab, Centro Investigación Príncipe Felipe (CIPF), València, Spain

Funding with a project PROMETEO 2023 (CIPROM/2022/25) of the Conselleria d'Educació, Universitats i Ocupació.

Introduction

Spinal cord injury (SCI) is a leading cause of paralysis, and currently, there is no cure capable of restoring lost neurological function in cases of complete lesions. However, partial recovery of locomotor function is sometimes possible in incomplete lesions, though the underlying mechanisms remain poorly understood. In this study, we employ **Visium Spatial Transcriptomics** (ST) and **single-cell RNA sequencing** (scRNA-Seq), combined with high-resolution imaging, to analyze **4 spinal cord samples** from 2 **healthy** rats and 2 rat 10 days post-spinal cord **injury** (SCI). The main objectives are to decipher the **biological processes** or molecular functions that are differentially represented in each of the anatomical regions of the spinal cord upon injury, as well as the distinct **cellular composition** in each area and the cell types that are responsible of the changes in gene expression.

Pipeline and Methodology



Cell2location vs RCTD/CSIDE

Cell2location

$$d_{s,g} \sim NB\left(\mu_{s,g}, \alpha_{e,g}\right) \qquad \mu_{s,g} = \left(\underbrace{m_g}_{\text{technology sensitivity}} \cdot \underbrace{\sum_f w_{s,f}\, g_{f,g}}_{\text{cell type contributions}} + \underbrace{s_{e,g}}_{\text{additive shift}} \right) \cdot \underbrace{y_s}_{\text{per-location sensitivity}}$$

Gene Expression per cell type

$$u_{s,f} = w_{s,f}\left(\sum_g m_g\, g_{f,g}\right)$$

RCTD + CSIDE

$$Y_{i,j,g} | \lambda_{i,j,g} \approx \text{Poisson}(N_{i,g}\lambda_{i,j,g}), \qquad \log(\lambda_{i,j,g}) = \log\left(\sum_{k=1}^{K} \beta_{i,k,g}\mu_{i,k,j,g}\right) + \gamma_{j,g} + \varepsilon_{i,j,g}, \qquad \log(\mu_{i,k,j,g}) = \alpha_{0,k,j,g} + \sum_{\ell=1}^{L} x_{i,\ell,g}\alpha_{\ell,k,j,g}.$$

	Input	Model type	Distribution	Output cell type Weights	Output cell type gene expression
Cell2location	ST + single cell (annotated CT)	Bayesian	Negative Binomial	Per spot	Per spot For celltype
RCTD + CSIDE	ST + single cell (annotated CT)	Probabilistic + linear mixed model	Poisson	Per spot	Per area For 1 cell

Cell types tested per area

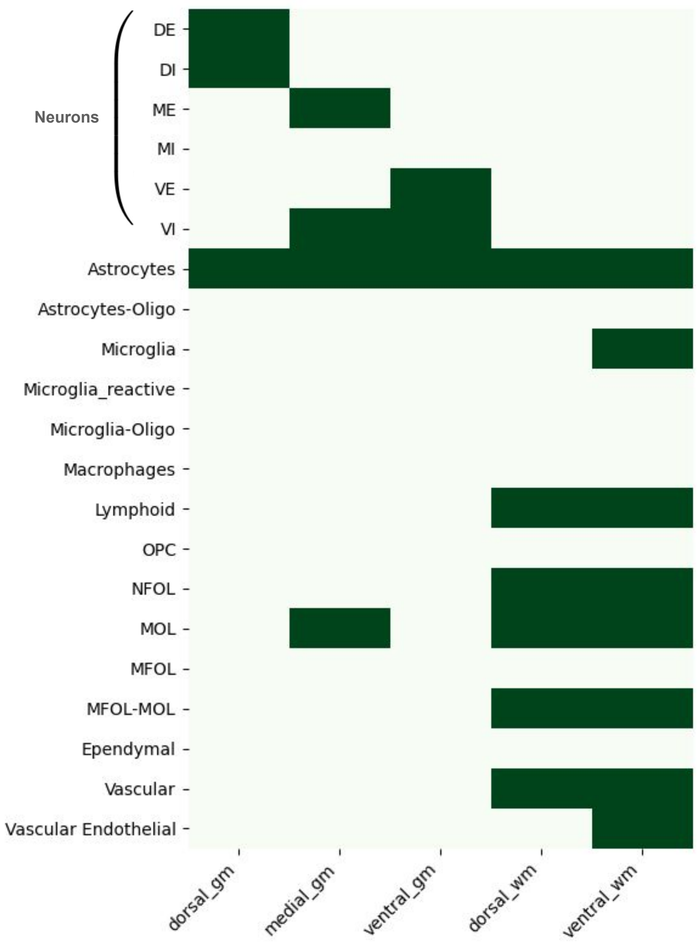
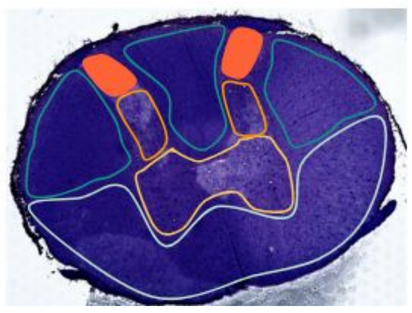
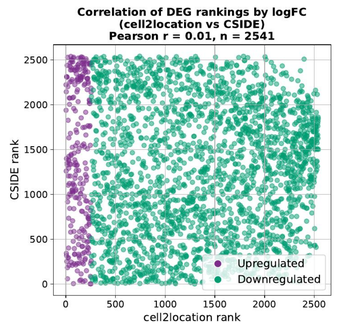
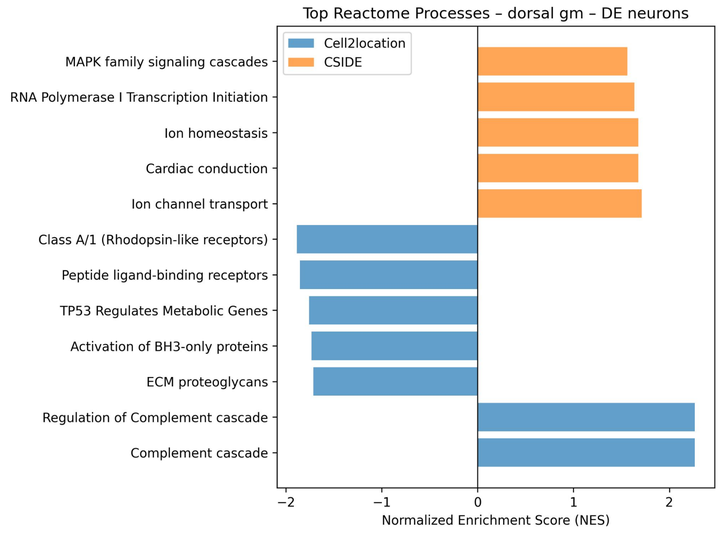


Figure 1. Cell types included for testing differential gene expression in each area.

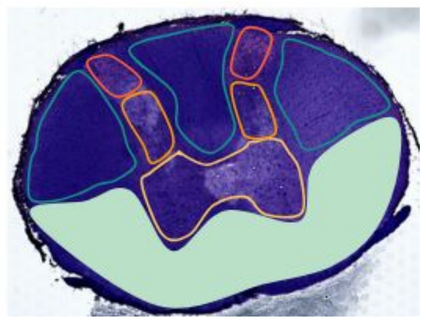
Dorsal Gray Matter - Dorsal Excitatory Neurons

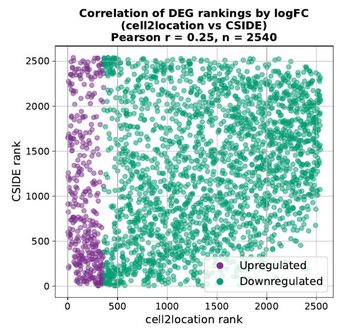






Ventral White Matter - Mature Oligodendrocytes





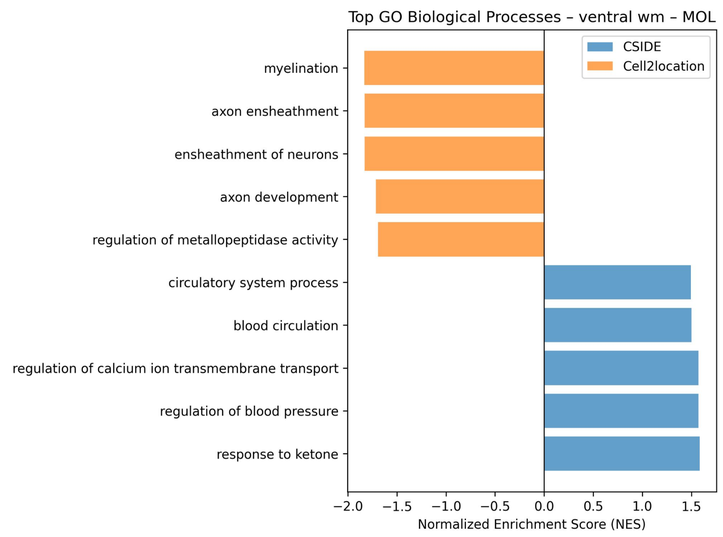


Figure 2. Comparison of differential expression results between cell2location and RCTD + CSIDE for 2 spinal cord regions and 2 cell types.

(A) Pearson correlation of ranked log fold changes (logFC) for the list of genes tested by both methods, split by direction.

(B) Top enriched processes for DE Neurons in the dorsal gray matter area (left) and for Mature Oligodendrocytes in the ventral white matter (right).

Conclusions

- **Cell2location** and **RCTD + CSIDE** both estimate **cell-type-specific gene expression** and identify **Differentially Expressed Genes (DEGs)** from ST data.
- The **ranking and magnitude of DEGs and enriched processes** differ markedly between both methods.
- Even after correcting Cell2location estimates for cell-type gene expression values to approximate per-cell values, discrepancies persist due to differences in **model assumptions, scale and inference targets.**
- Together, they provide **complementary views** of spatial gene regulation.