

Ruoxi Zhao¹, Margot Bezruczyk², Karen Serrano³, Henrik Scheller³, Benjamin Cole²
¹University of California, Merced, ²Lawrence Berkeley National Laboratory, ³Joint Bioenergy Institute

ABSTRACT

We are interested in understanding how plants like *Medicago* interact with symbiotic fungi. To do this, we are integrating different types of single cell datasets to build a comprehensive map.

BACKGROUND INFO

Plant roots act not only as a mechanism to get water from the soil, but also as an organ to interact with the microbe and fungi. In order for the plants to grow, we need key nutrients such as nitrogen, phosphorus, and potassium. This is important because we are running out some types of fertilizers, due to the finite amount of rock phosphate, and nitrogen is fossil-fuel intensive to produce. We want to study the plant-microbe interaction for sustainable biological sources of these nutrients.

Medicago and AMF

Arbuscular Mycorrhizal Fungi (AMF) supplies *Medicago* with phosphorus.¹ *Medicago* is a good model organism, it has a well-categorized genome. The strain of rhizophagus we are using has a recently sequenced genome and is commonly used in commercial agriculture. Because *Medicago* and *Rhizophagus* are both eukaryotes, we can use the same poly(A) tail capture mechanism to obtain mRNA from both organisms in one assay.

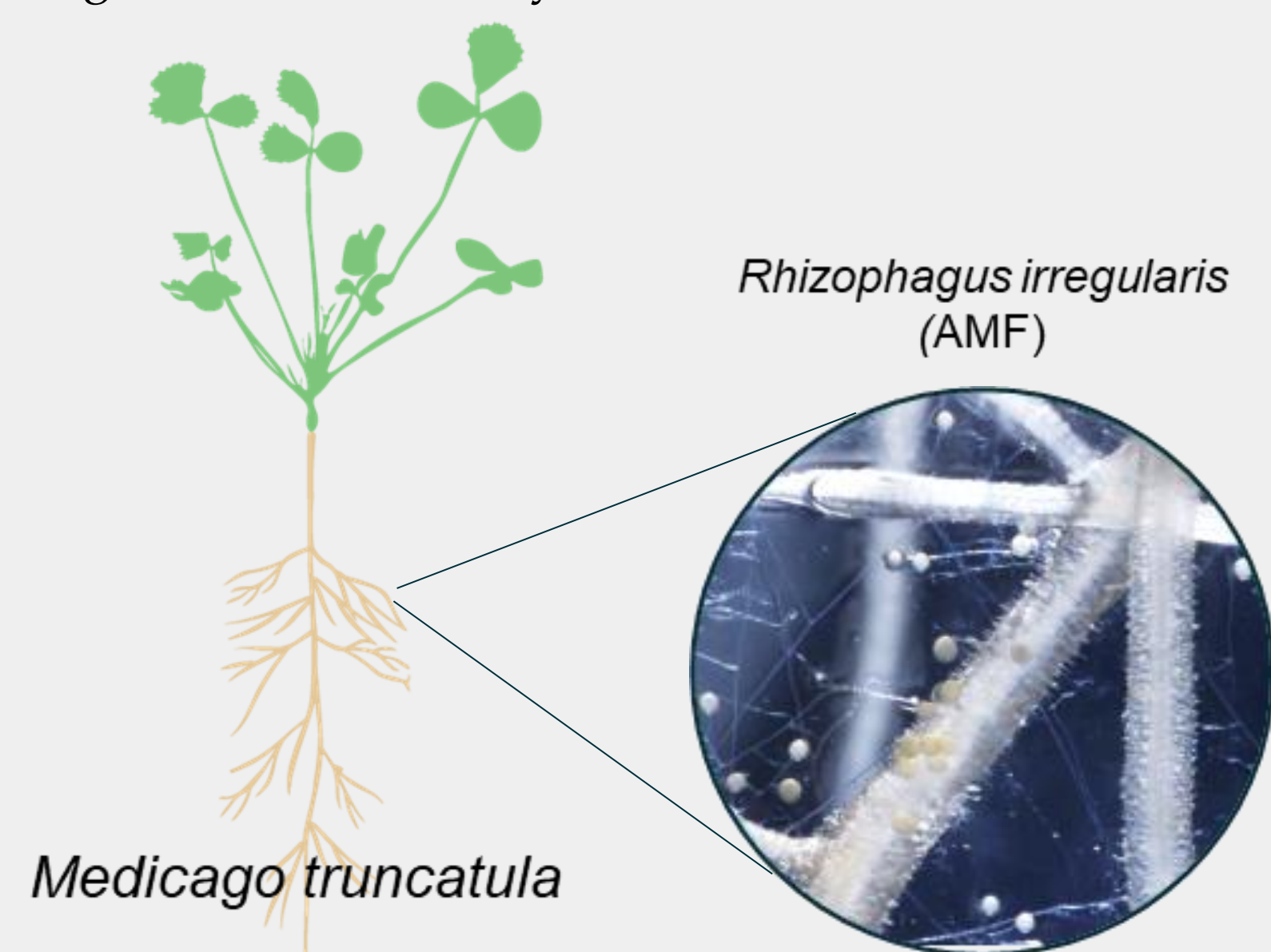


Figure 1. Plant roots and Arbuscular Mycorrhizal Fungi symbiosis. (a) *Medicago truncatula* and *Rhizophagus irregularis* (AMF), Margot Bezruczyk et al (unpublished) (b).

Tangram

Tangram⁴ is a Python package, written in PyTorch and based on scanpy. Tangram⁴ learns a spatial alignment of the single cell data so that the gene expression of the aligned single cell data is as similar as possible to that of the spatial data. Using integration, Tangram⁴ creates new spatial data by aligning the scRNAseq profiles in space using cosine similarity. Tangram⁴ uses a cosine similarity loss function to calculate the gene expression of the mapped single cells being as similar as possible to the spatial data.⁵

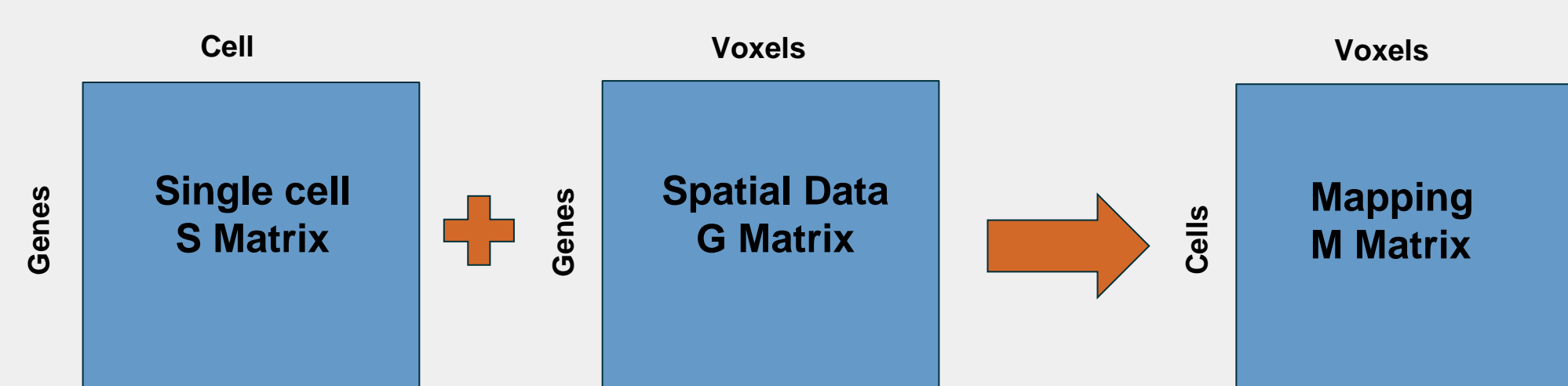


Figure 2. Simplified algorithm of Tangram⁴ integrating the single-cell dataset and spatial dataset to mapping dataset as matrices

Spatial Transcriptomics

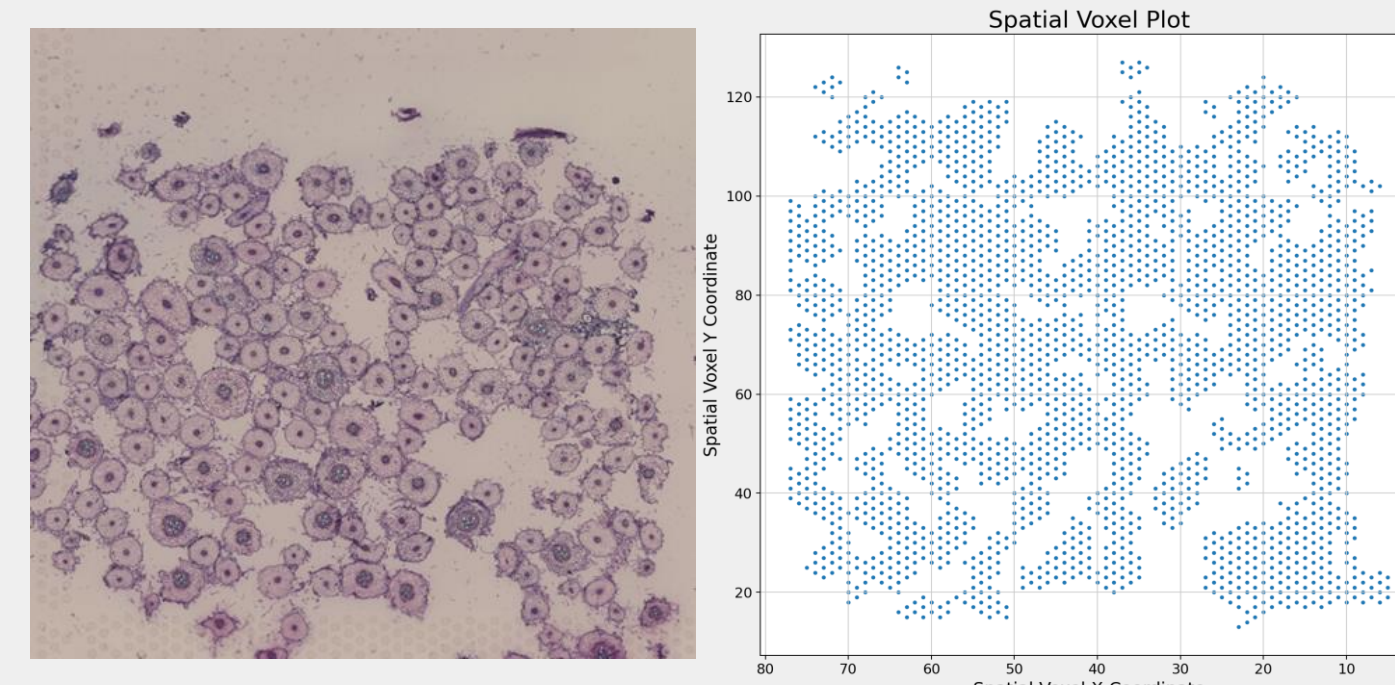


Figure 3. 3.16µm transverse sections of *Medicago* root tissue colonized with AMF, Margot Bezruczyk et al (unpublished) (a). Voxels representing xyz (b).

The totality of the mRNA in the cell is the transcriptome, which tells you the current state of the cell activity and conveys what protein to translate. Spatial transcriptomics is a molecular profiling method that we use to measure the amount of mRNA present for each gene in the tissue sample, and map where the activity is occurring.²

Each root has various types of cells (a), and each spot in the plot represents one voxel that contains multiple cells (b).

Single Cell RNA Seq

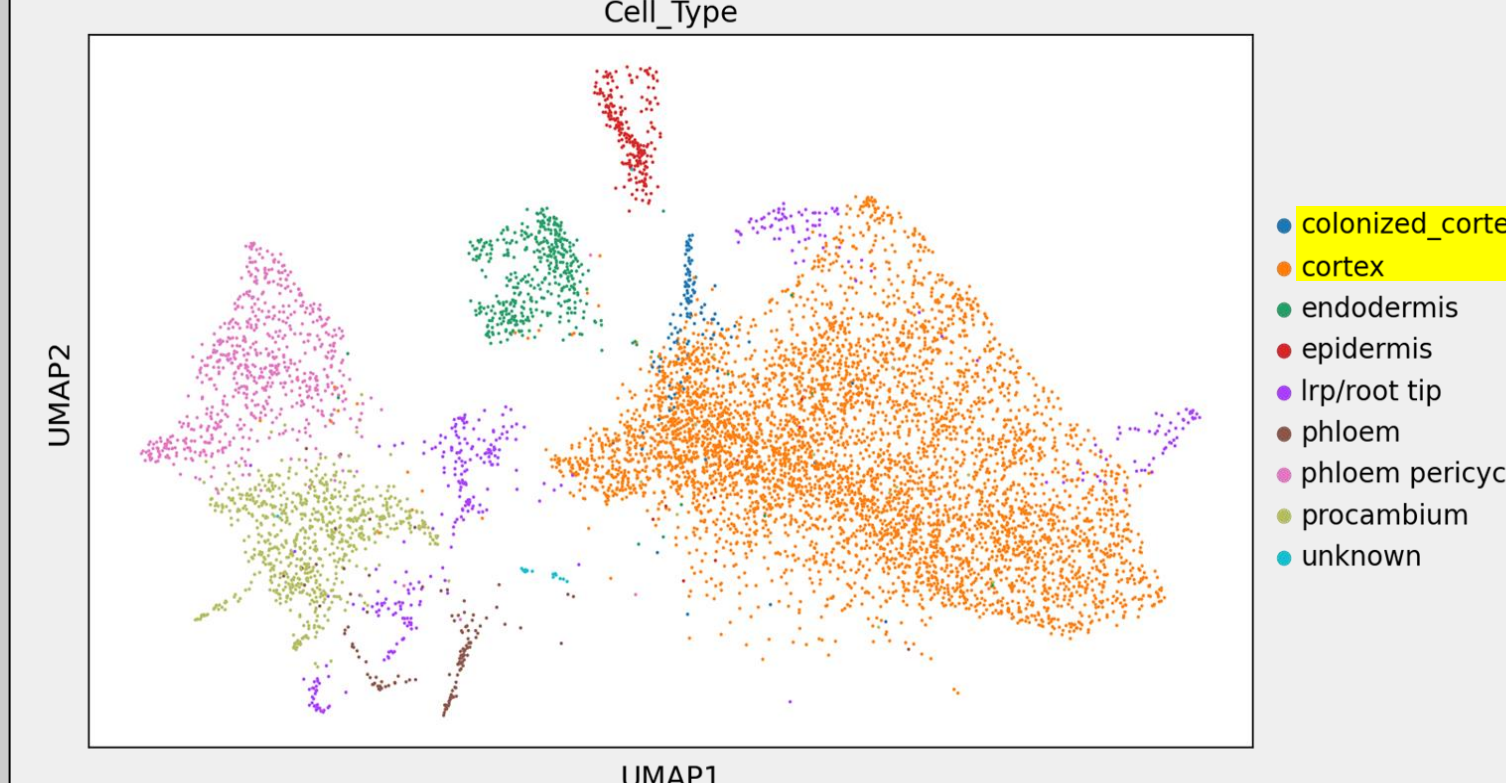


Figure 4. Single-cell dataset of chromium (single cell), created using the 10X Genomics system

Another technique we are using to understand this system is single cell sequencing.

The dots represent the transcriptome of each cell, and their position in 2D UMAP space is related to the similarity of transcriptome profiles. The most common cell type in this dataset is the cortical cell. Cortex is the tissue of unspecialized cells lying between the epidermis (surface cells) and the vascular, or conducting, tissues of stems and roots.³

Cell Types mapped onto Space

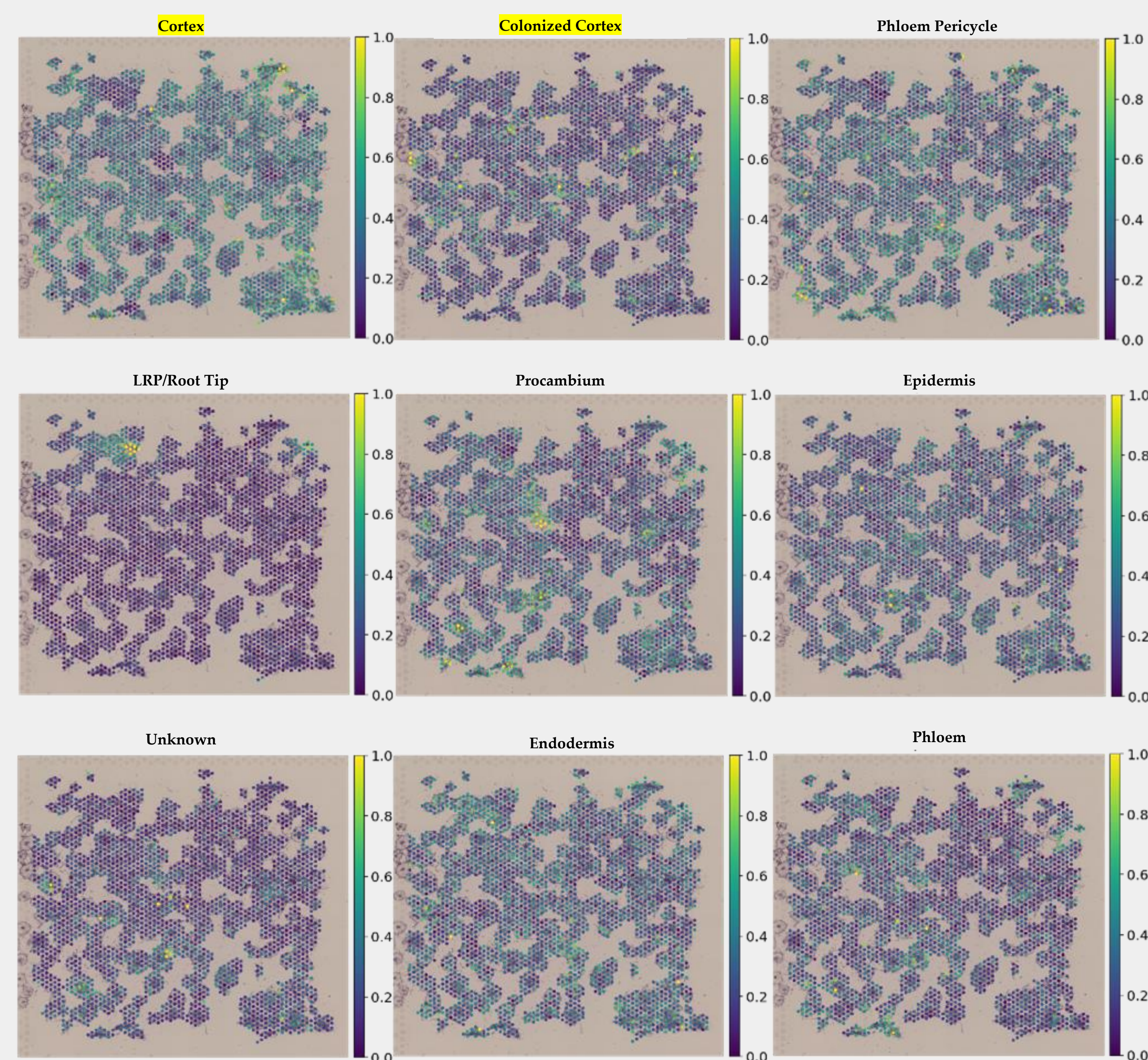


Figure 5. Mapping cell type annotations onto space using Tangram⁴.

Heat maps representing the predicted locations of different cell types in the tissue section using Tangram⁴. We want to focus on the predicted location of colonized cortex (yellow box) to study this plant-microbe interaction.

Progress and Results

- Load AnnData objects of the single cell data and spatial data in NERSC Jupyter Notebook.
- Clean and normalize the datasets
- Plot single cell data in UMAP and spatial data with voxels in xyz coordinates.
- Use PyTorch to fit the model with training genes. The score measures the similarity between the gene expression of the mapped cells and spatial data
- Cell mode in Tangram⁴ maps the data at single cell level (a). In cluster mode, we average the cells within a cluster and map the averaged cells instead. Cell mode takes longer to map but is more accurate. (b)

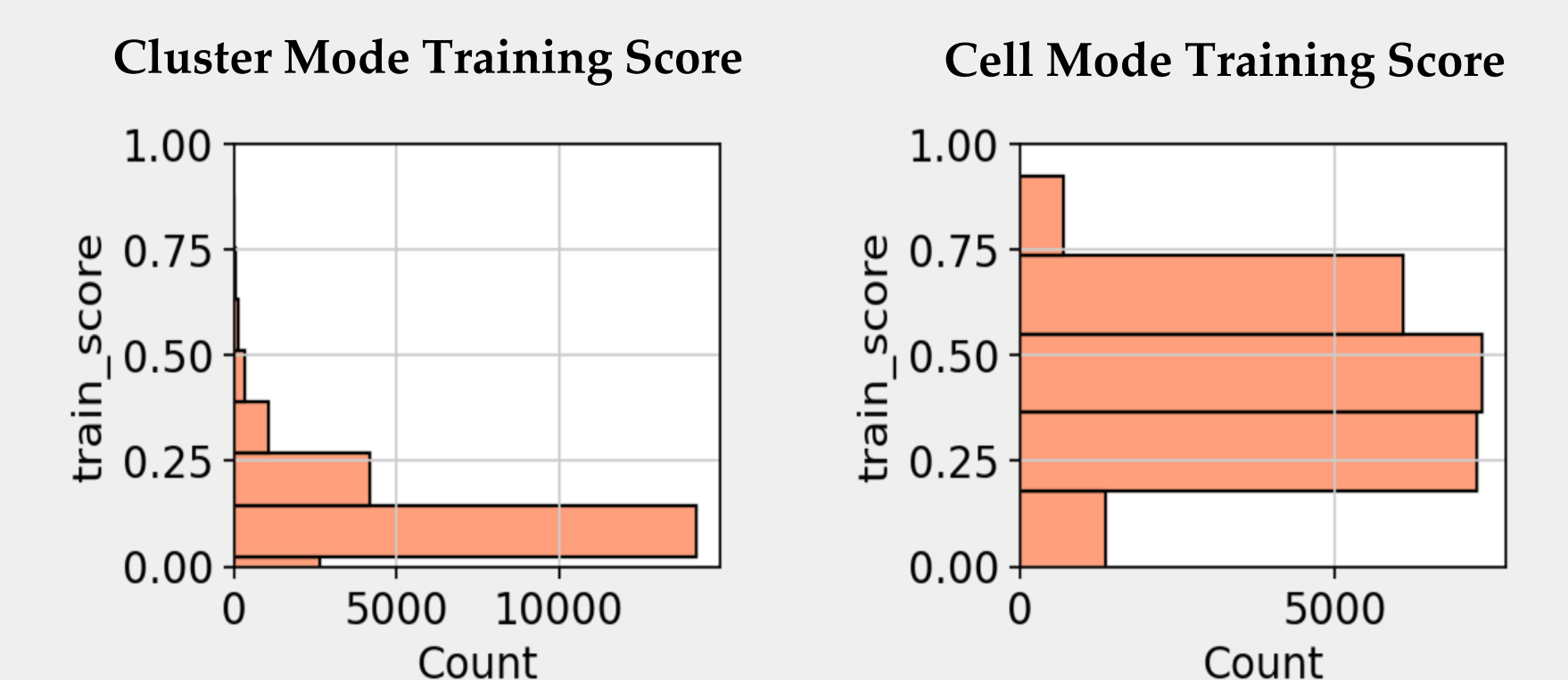


Figure 6. Tangram⁴ training score with cluster mode. (a) Tangram⁴ training score with cell mode. (b)

Future Work

- Integrate and visualize both single cell data and spatial transcriptomics on one plot.
- Decompose the voxels into single cells using cell segmentation.

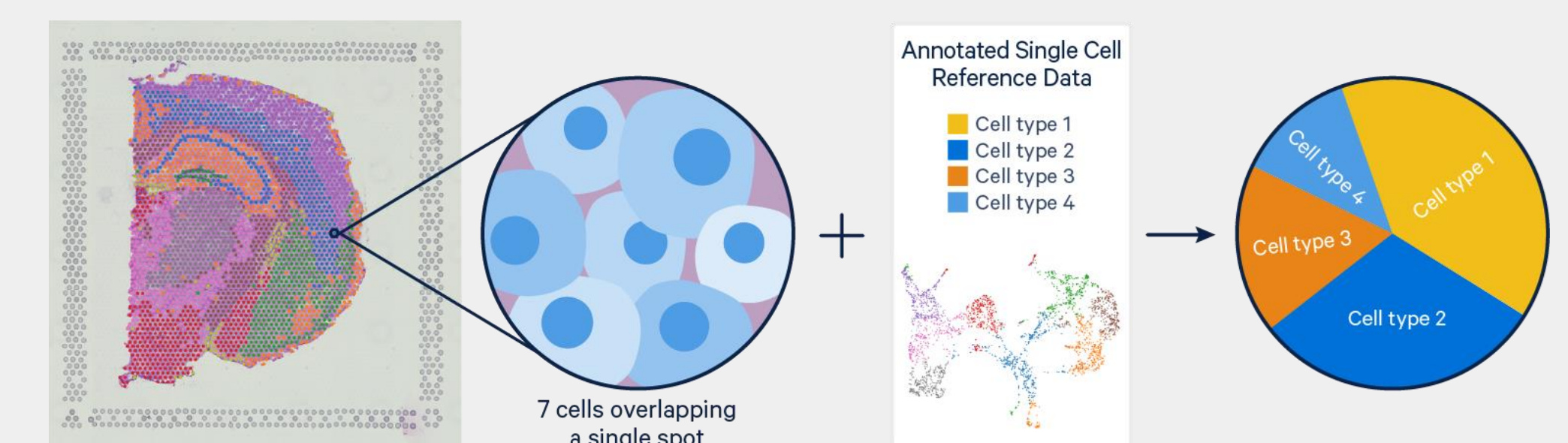


Figure 7. Overlapped cells in one voxel and annotated single-cell types, 10X Genomics⁶

Cell Segmentation using Cellpose

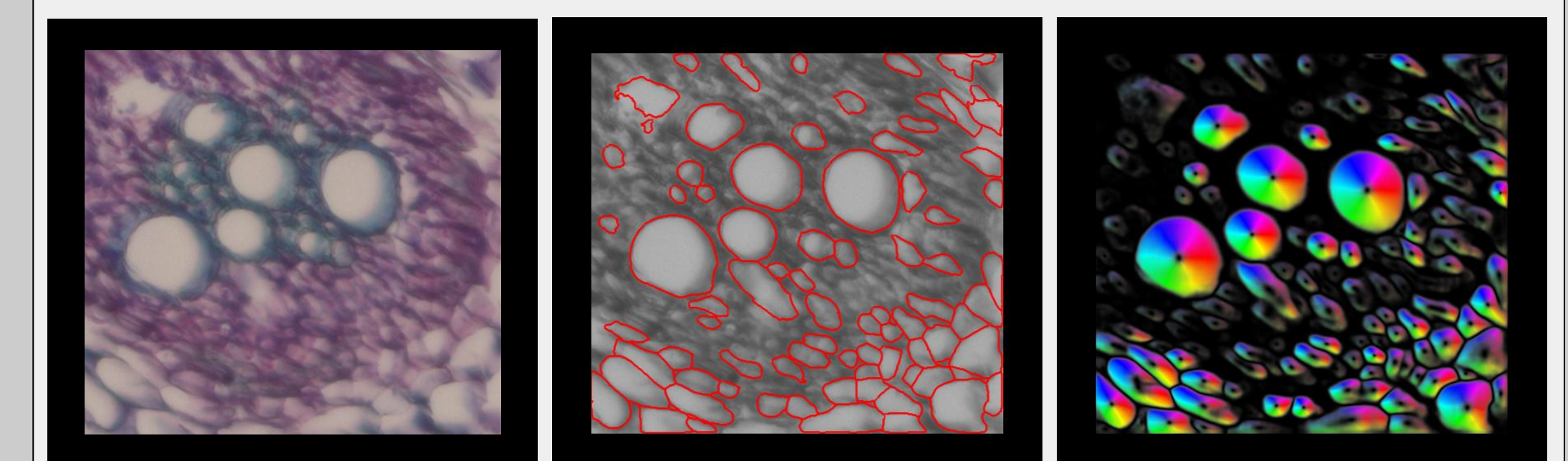


Figure 8. Cropped image of chopped *Medicago* root section (a), Cellpose⁷ predicted outlines (b), predicted cell pose (c)

References

- ¹Parishko, M. Arbuscular mycorrhiza: the mother of plant root endosymbiosis. *Nat. Rev. Microbiol.* 6, 763–775 (2008). <https://doi.org/10.1038/nrmicro1987>
- ²10X Genomics. Spatial Transcriptomics. <https://www.10xgenomics.com/spatial-transcriptomics>
- ³Britannica. The Editors of Encyclopedia Britannica. "cortex." *Encyclopedia Britannica*, 23 Nov. 2011. <https://www.britannica.com/science/cortex-plant-tissue>.
- ⁴Bhargava, T., Scallia, G., Buffoni, L. et al. Tangram GitHub. <https://github.com/broadinstitute/Tangram>
- ⁵Bhargava, T., Scallia, G., Buffoni, L. et al. Deep-learning and alignment of spatially resolved single-cell transcriptomes with Tangram. *Nat. Methods* 18, 1352–1362 (2021). <https://doi.org/10.1038/s41592-021-01264-7>
- ⁶10X Genomics. Integrating Single-cell and Visium Spatial Gene Expression Data, 16 March, 2022. <https://www.10xgenomics.com/resources/analysis-guides/integrating-single-cell-and-visium-spatial-gene-expression-data>
- ⁷Carsen Stringer, Marius Pachitariu bioRxiv 2022.04.01.486764; doi: <https://doi.org/10.1101/2022.04.01.486764>; <https://github.com/MouseLand/cellpose/tree/main>
- Alyson M. MacLean, L. Armando Bravo, and Maria J. Harrison. Plant Signaling and Metabolic Pathways Enabling Arbuscular Mycorrhizal Symbiosis

ACKNOWLEDGMENTS

Many thanks to my mentor Margot Bezruczyk, Benjamin Cole, Axel Visel, Jonelle Basso, Natasha Brown, Marlyn Hui, Jorge Arroyo, the Plant-Microbe group, JGI, Berkeley Lab and UC Merced. Thank the Department of Energy's Workforce Development of Teachers and Scientists as well as Workforce Development & Education at Berkeley Lab.

