# Integrating Gene Expression in the Alignment of Spatial Transcriptomics Data

OSAGIE K. AIMIUWU[1,2], Kalen Clifton[3], Michael I. Miller[3], Jean Fan[3], Daniel Tward[4]

[1]BIG Summer Program, Institute for Quantitative and Computational Biosciences, UCLA
[2]University of North Carolina, Chapel Hill
[3]Johns Hopkins University
[4]Department of Computational Medicine, Department of Neurology, UCLA

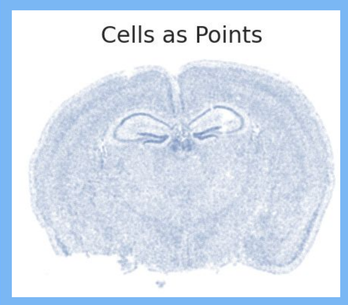
## Background/Motivation

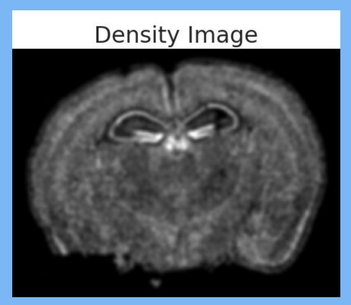
Spatial transcriptomics (ST) technologies like MERFISH can reveal novel information about cell types by providing the full gene expression profiles of cells in tissues, including the brain.

The ability to align, or register, two brain sections could potentially lead to a more systematic approach when comparing brain structures and regions between sections.

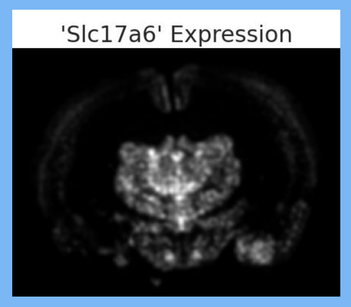
The aim of this study is to investigate how gene expression information should be included during the alignment procedure. This was accomplished by utilizing ST datasets consisting of three coronal brain sections with three replicates each.

## Methods/Results



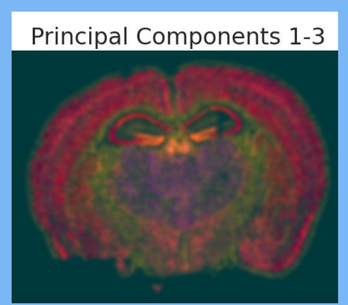



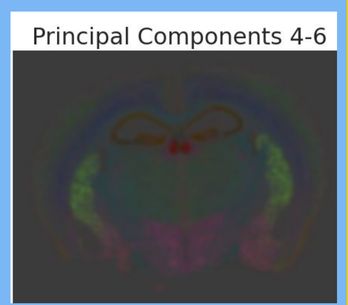


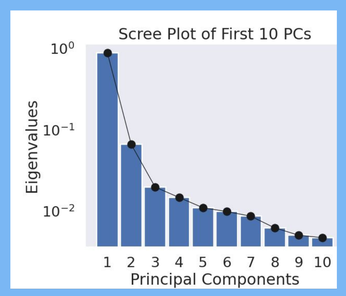


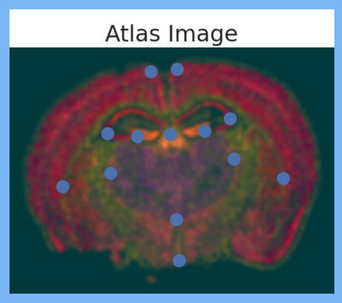

**1. Raw Data.** Images corresponding to the ST dataset for Slice 2, Replicate 3 (S2R3). The rasterized images were created using a Gaussian kernel. **Top:** Cell locations as a scatter plot **(left)** and a rasterized image **(right)**. **Bottom:** Rasterized images with specified gene expression info.

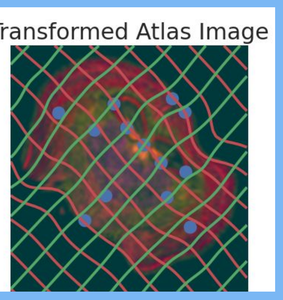
**2.PCA.** S2R3 after applying principal component analysis (PCA) to the 483 rasterized images that included gene expression information. **Top:** Images with PCs 1, 2, and 3 or 4, 5, and 6 in the red, green, and blue color channels, respectively. **Bottom:** Scree plot on a log scale.

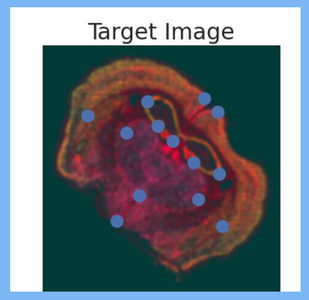


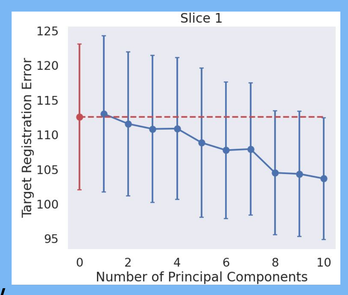



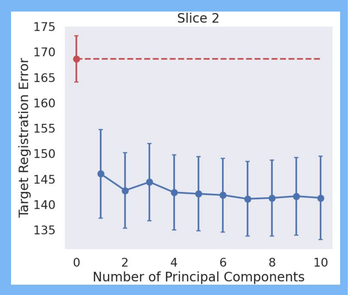



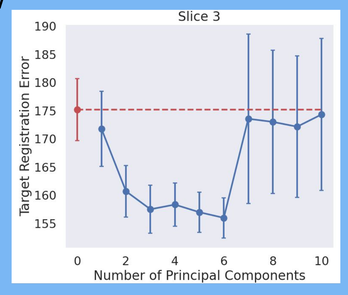



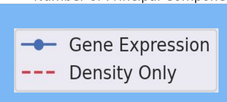

**3. Registration.** Alignment of S2R3 (atlas) to S2R2 (target) using manually annotated points as landmarks. **Top:** Atlas image with landmarks. **Bottom left:** Atlas image after applying affine and diffeomorphic transformations. Landmarks and coordinate grid lines are shown. **Bottom right:** Target image with landmarks.

**4. Accuracy.** Target registration error (TRE) of the alignment pairs as a function of the number of PCs included in the registration. The plotted points show the average TRE for a pair's landmarks, averaged across the six total pairs. Error bars are the standard error of the mean.




**Note:** Since PCA was not performed on the density images, there is only one TRE value per slice (shown in red).

## Conclusions

We found that the alignment with gene expression generally had higher accuracy than the density images

More PCs tend to increase accuracy and more work is required to identify an optimal number of PCs

Next steps include quantifying alignment accuracy at a more fine-grain level and building a software package to share methods with the scientific community

## Acknowledgements

National Institute of Mental Health: U19MH114821 and RF1MH128875
The Karen Toffler Charitable Trust
BIG Summer Program