

# Squidpy: Spatial Single Cell Analysis in Python

Giovanni Palla, SCOG workshop, 23rd May 2022



@giovp   
@g\_palla1 

# Lack of spatial single cell analysis framework in the python ecosystem

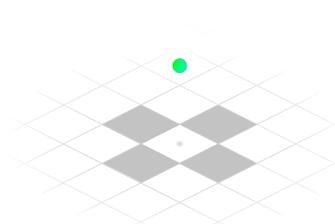
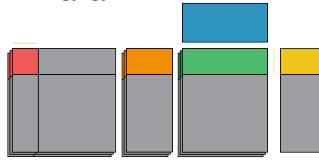


Single cell analysis in python

*Wolf, Angerer, Theis 2018*

- > 1000 citations
- ~ 400k downloads
- > 80 contributors

*AnnData*



Muon: multimodal omics

Python framework

*Bredikhin et al. 2022*



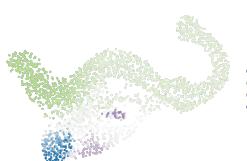
Probabilistic models  
for single-cell omics  
data

*Gayoso et al. 2021*



Scirpy: single-cell  
immune-cell receptor  
sequencing tool

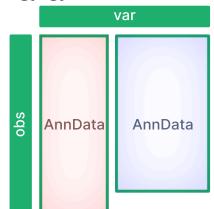
*Sturm et al. 2020*



CellRank - Probabilistic Fate  
Mapping using RNA Velocity

*Lange et al. 2020*

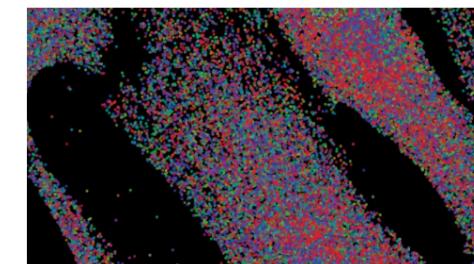
*MuData*



**nature methods**

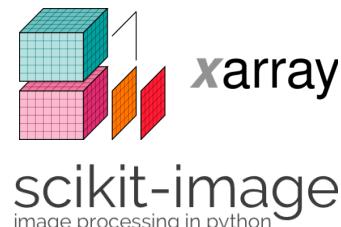
FOCUS | 06 JANUARY 2021

**Method of the Year 2020:  
spatially resolved  
transcriptomics**



**scverse**

Foundational tools for omics data in the life sciences



**xarray**

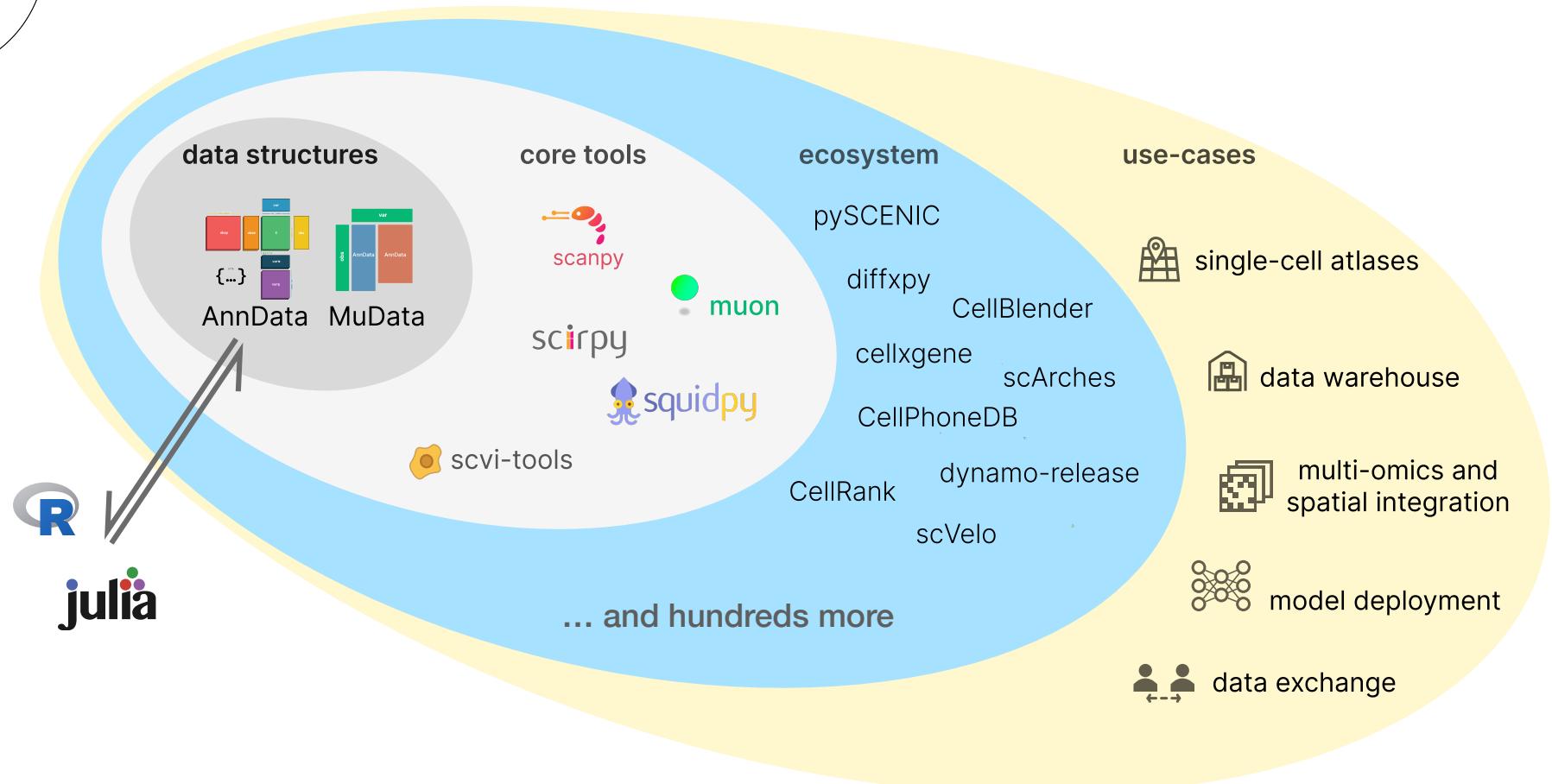


**Numba**

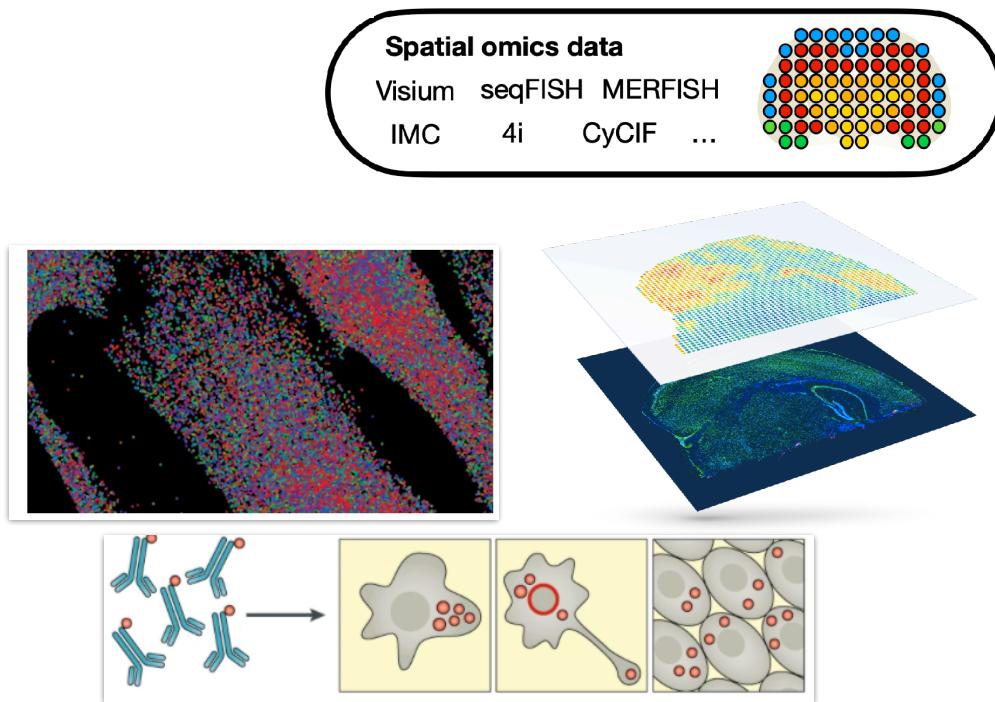


# scverse

Foundational tools for omics data in the life sciences



# Spatial data are diverse and require specific preprocessing methods



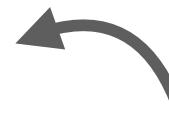
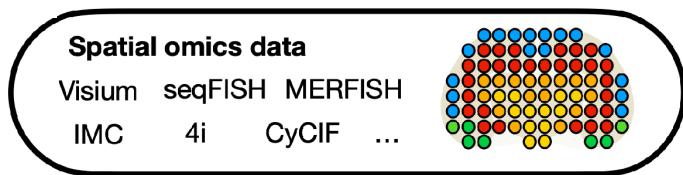
Larsson et al. 2021  
Zhuang et al. 2021  
Close et al. 2021  
Asp et al. 2020  
Lundberg et al. 2019  
10x Genomics

## Preprocessing tools

- *IMCtools* (*Bodenmiller group*)
- *Cycif processing* (*Sorger lab*)
- *seqFISH-PLUS* (*Cai group*)
- *MERFISH* (*Zhuang lab*)
- *Spaceranger* (*10x genomics*)
- *Starfish* (*CZI*)



# Squidpy provides data representations and downstream analysis tools



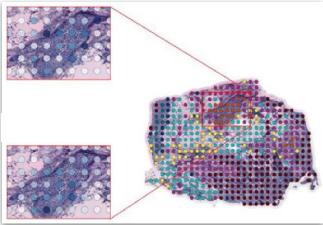
## Preprocessing tools

- *IMCtools* (*Bodenmiller group*)
- *Cycif processing* (*Sorger lab*)
- *seqFISH-PLUS* (*Cai group*)
- *MERFISH* (*Zhuang lab*)
- *Spaceranger* (*10x genomics*)
- *Starfish* (*CZI*)

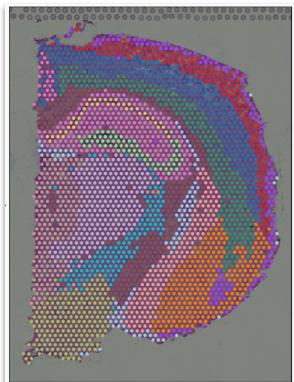


# Flexible spatial graph building tailored to various spatial technologies

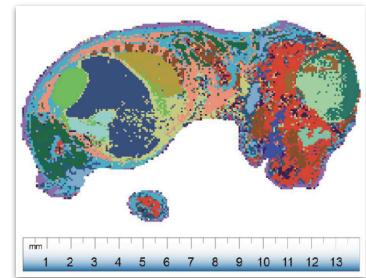
ST (Andersson et al. 2020)



Stereo-seq (Chen et al. BGI 2020)

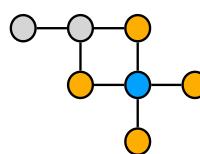


Visium (10x Genomics 2020)

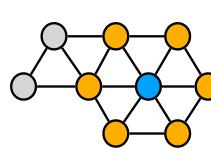


Grid

Square

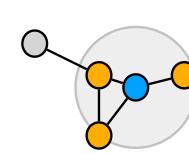


Hex

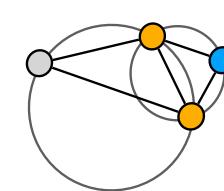


Generic

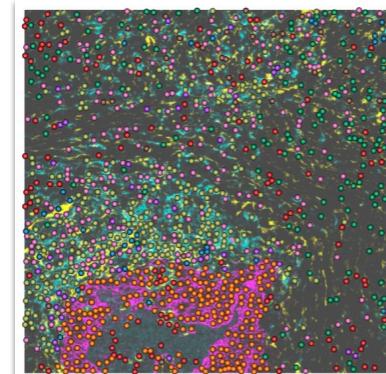
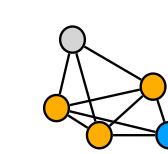
Radius



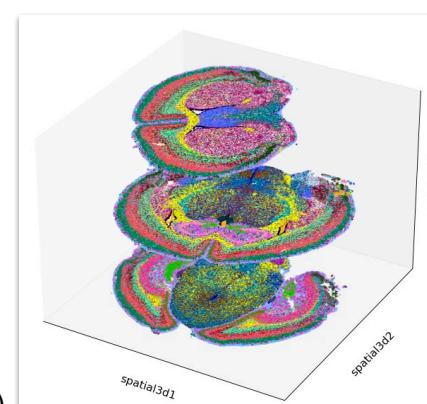
Delaunay



KNN

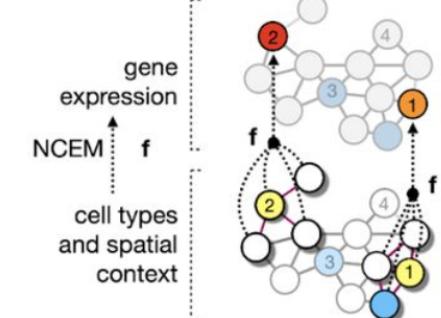


MIBI-TOF (Hartmann et al. 2021)



MERFISH (Vizgen 2021)

seqFISH+ (Lohoff et al. 2020)



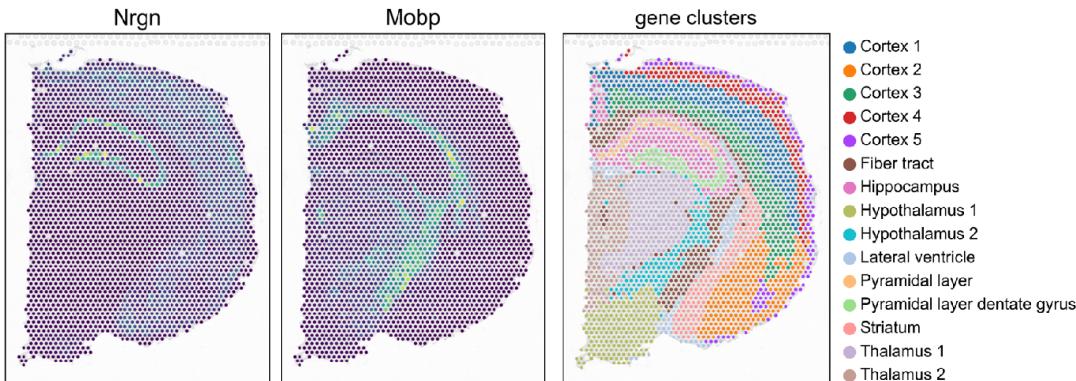
Node-centric expression models  
Fischer, Schaar, Theis  
Bioarxiv 2021

# Tools to extract spatial patterns in tissue

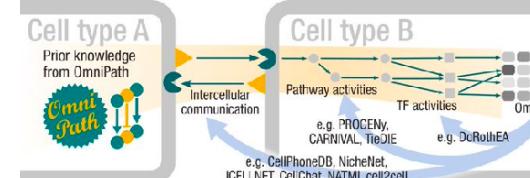
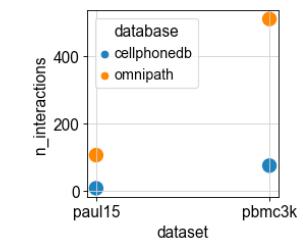
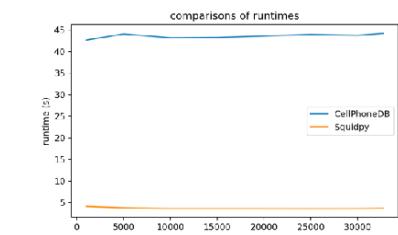
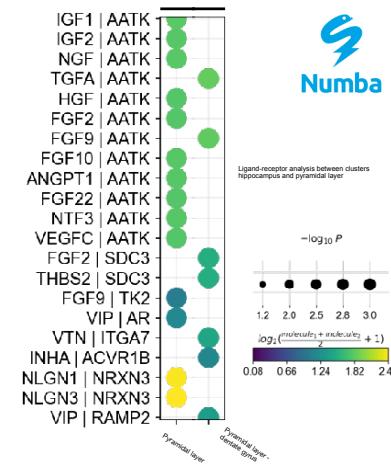
## Continuous features (e.g. genes)

- Moran's I spatial autocorrelation statistic (and orthogonal Geary's C)
- Sepal score (Andersson et al. 2021)
- CellphoneDB LR interaction test with Omnipath database

## Spatially variable genes: Moran's I spatial autocorrelation



## Ligand-receptor interaction: cellphoneDB



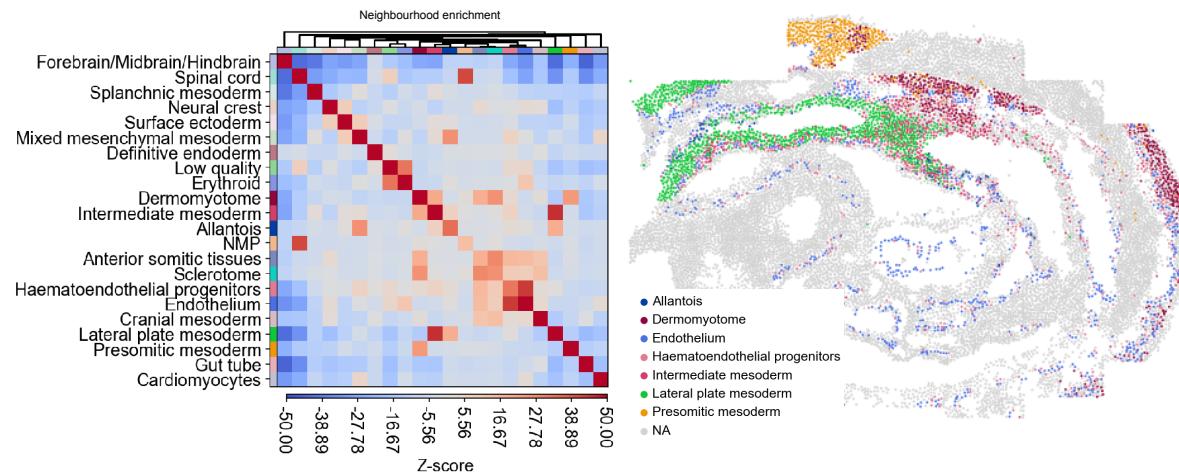
Efremova et al. 2020  
Turei et al. 2021

# Tools to extract spatial patterns in tissue

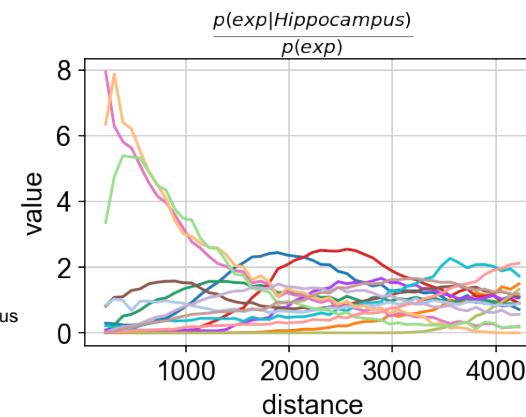
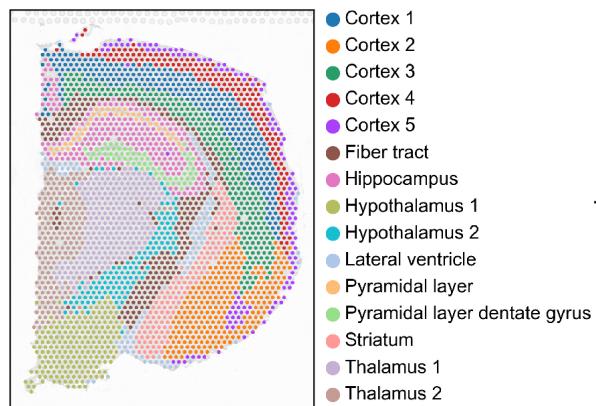
## Discrete annotations (e.g. clusters)

- Neighbourhood enrichment score
- Ripley's statistics
- Network centralities
- Co-occurrence score

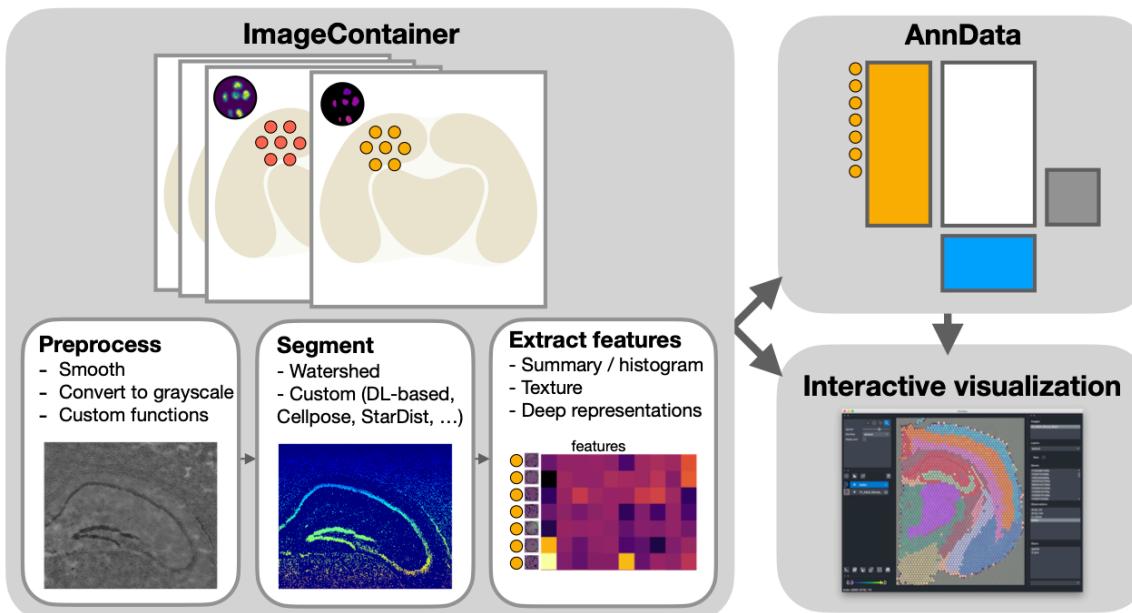
## Neighbourhood enrichment analysis



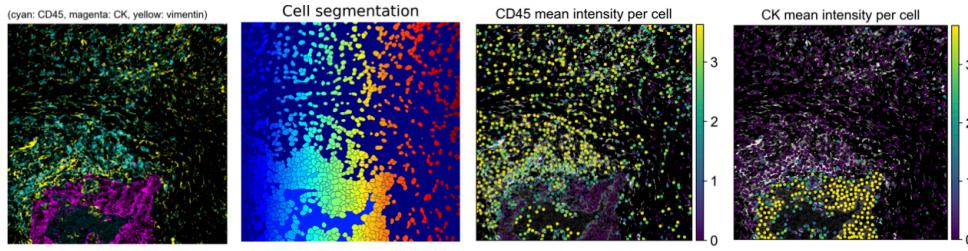
## Co-occurrence score for clusters



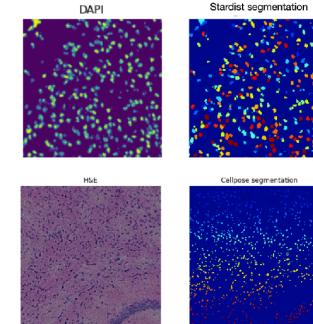
# The Image Container: integrating high-dimensional images with molecular data



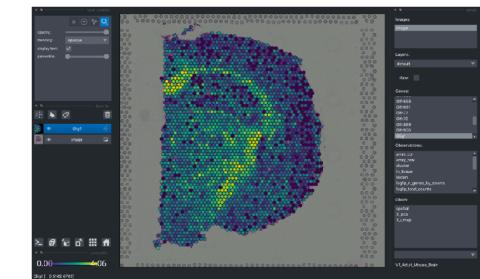
MIBI-TOF processing and visualization



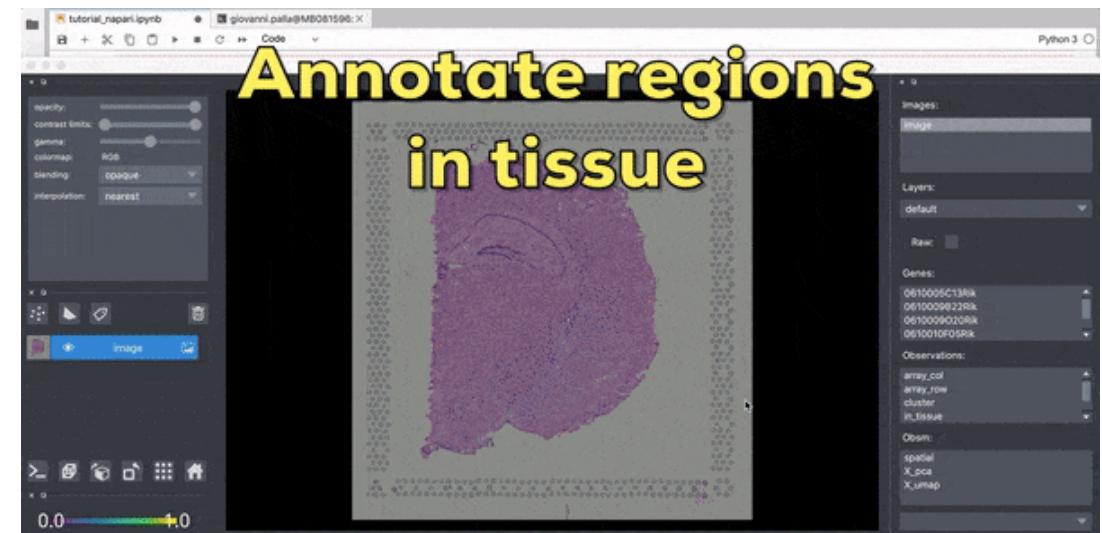
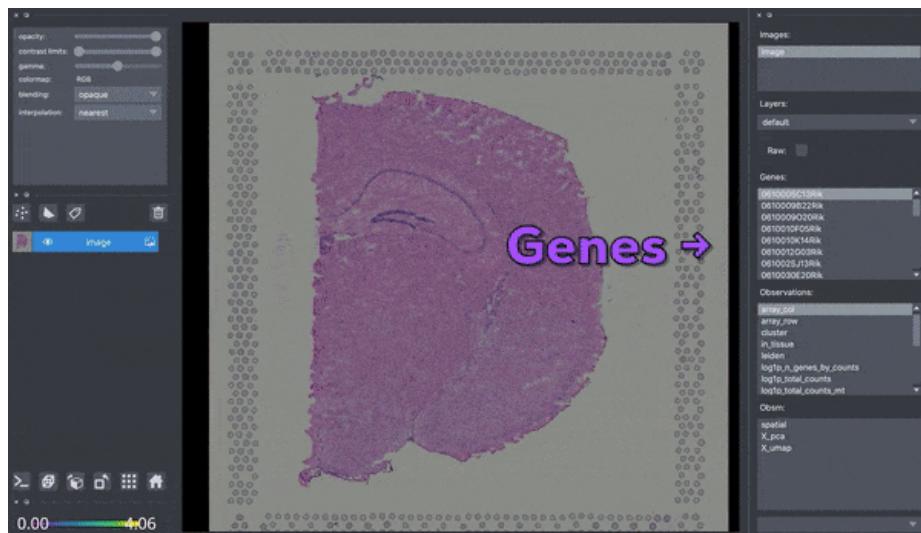
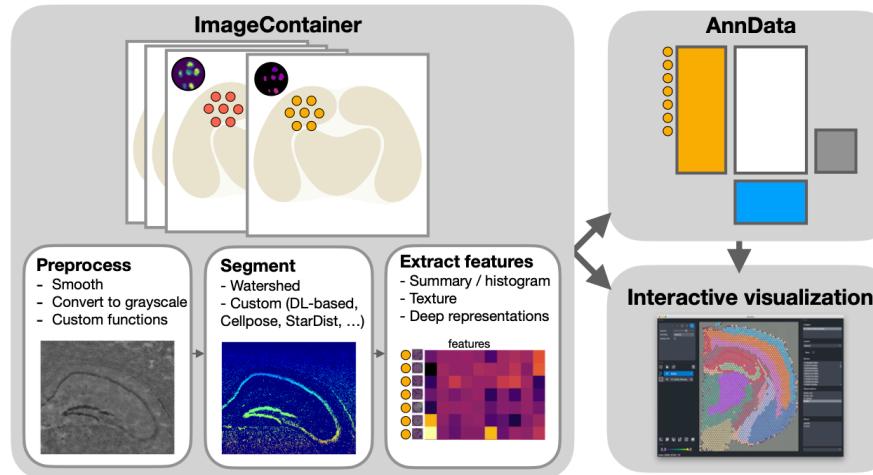
External tools for segmentation



Napari integration



# Napari plugin for interactive visualization of spatial omics data



# Squidpy 1.2 new features

195 stars, 34 forks, 14k downloads on PyPI

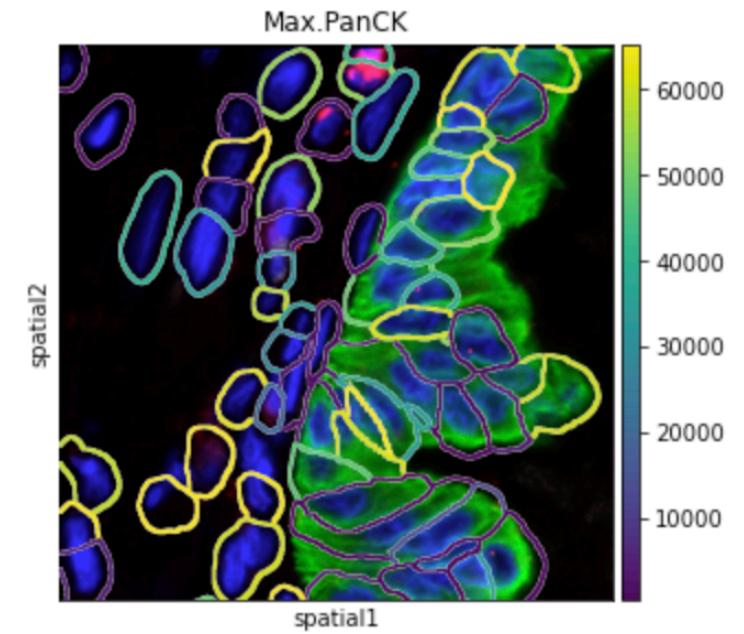
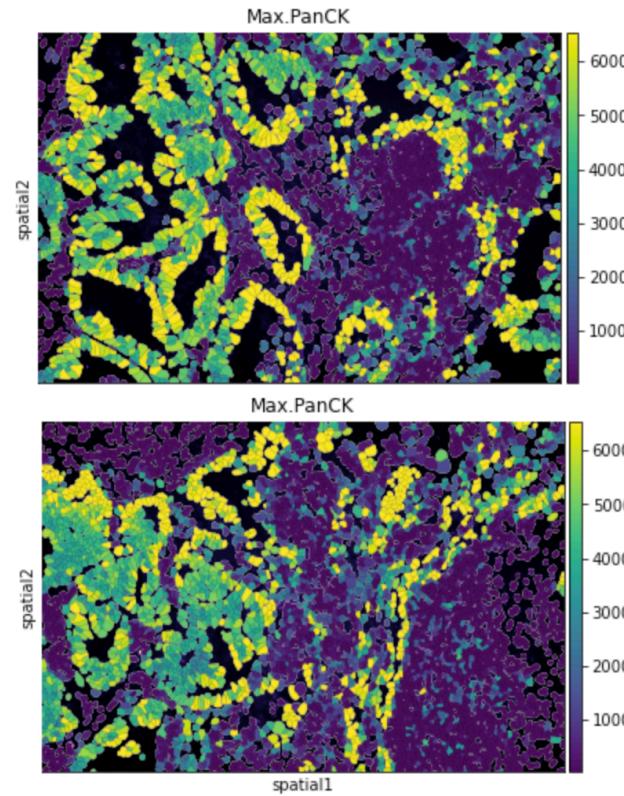
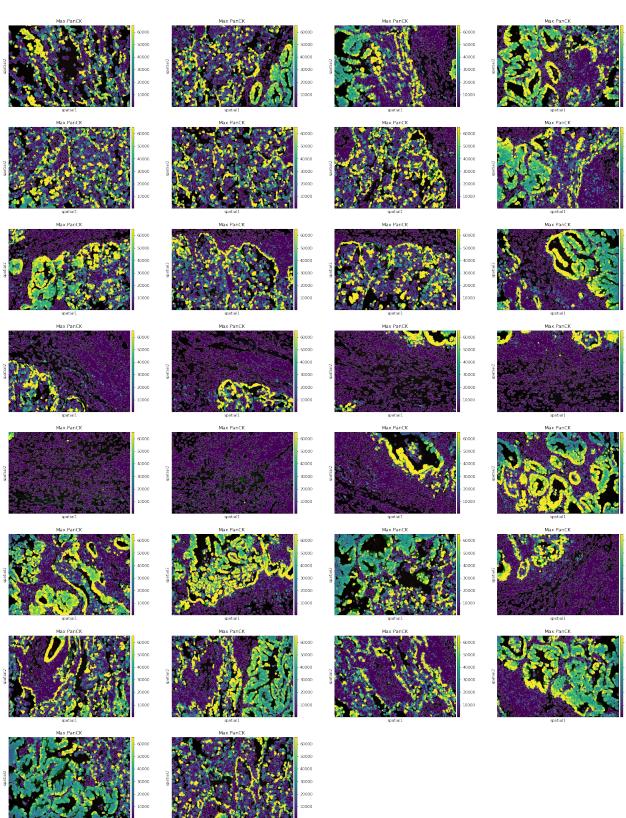
## New features:

- **Plotting:** static plotting for scatter, polygons and segmentation masks.
- **Datasets:** all 10x genomics datasets can be downloaded with 1 lines: `sq.datasets.visium()`
- **Graph building:** now possible to build spatial neighbours across slides.
- **Reading functions:**
  - `sq.read.nanostring`
  - `sq.read.vizgen`
  - `sq.read.visium`

+ *bug fixes and various additions*

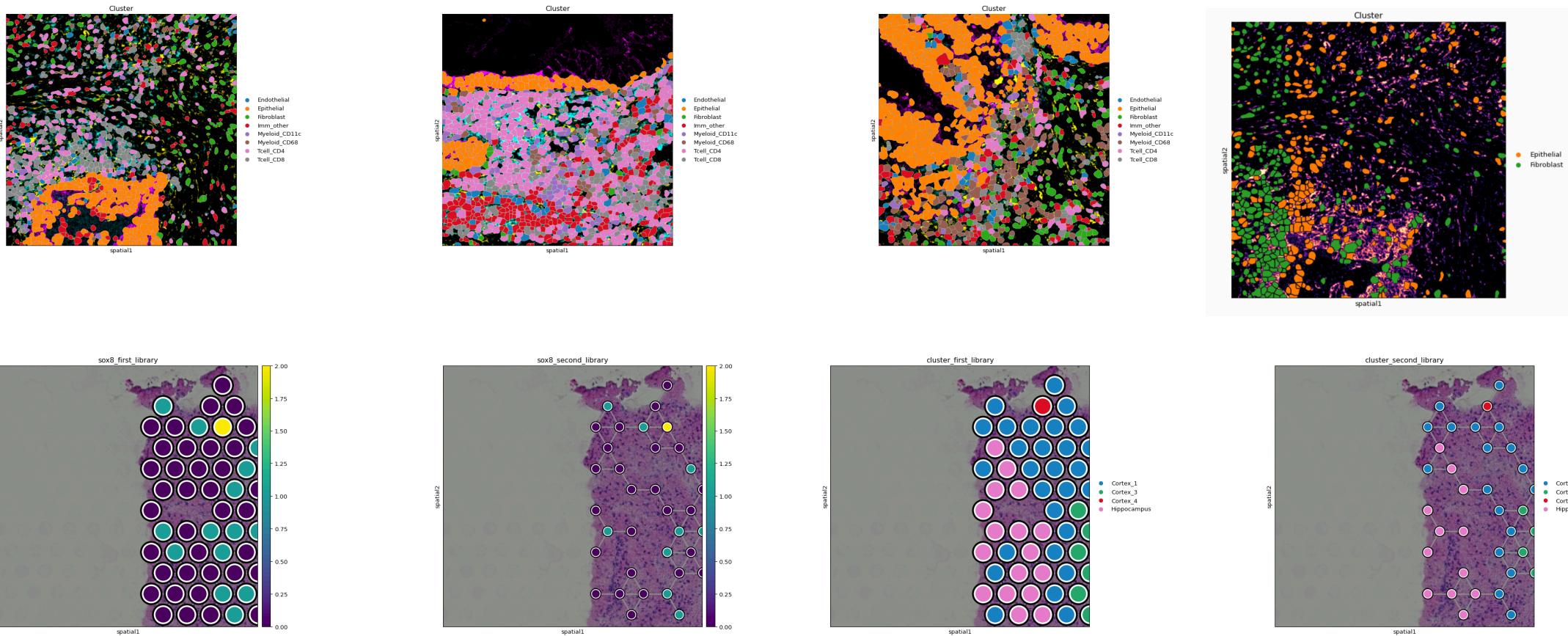
# Squidpy 1.2 example on nanostring data

```
sq.pl.spatial_segment(adata, color="Max.PanCK", library_key="fov", seg_cell_id="cell_ID",)
```



# Refactor and improved static plotting functionality in Squidpy

## Support for multiple-slides dataset, segmentation masks and polygons



# Future: spatial omics analysis in scverse

## Data formats

- New spatial data formats based on OME-NGFF + ZARR + ANNDATA  
with:

- Isaac Virshup (Helmholtz)
- Luca Marconato (EMBL)
- Kevin Yamauchi (ETH)
- + OME and scverse developers

## Data representation

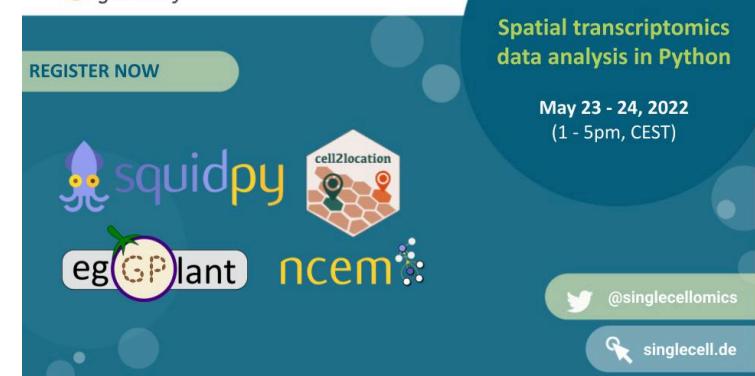
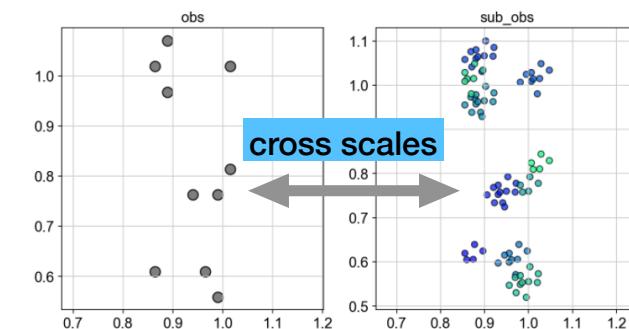
- Awkward arrays for storing and analysing transcript locations

## Workshops

- Single cell omics Germany next week (registration still available) at [www.singlecell.de](http://www.singlecell.de)

With:

- David Fischer, Anna Schaar (Helmholtz)
- Vitalii Kleshchevnikov (Sanger)
- Alma Andersson (KTH)



# Resources + acknowledgments

Docs: [squidpy.readthedocs.io](https://squidpy.readthedocs.io)

Code: [github.com/theislab/squidpy](https://github.com/theislab/squidpy)

Paper: Palla, Spitzer et al. Nature Methods 2022



A screenshot of the squidpy documentation website. The main navigation bar includes links for Home, API, and GitHub. The left sidebar has sections for General, Data types, Graph, Tutorials, Examples, Core, and Galleries. The main content area shows a "Tutorials" section with a sub-section for "Graph". Below it are examples for "Compute centrality scores", "Compute Co-occurrence probability", "Compute interaction matrix", and "Receptor-ligand analysis".

MUDS MUNICH SCHOOL FOR  
DATA SCIENCE  
HELMHOLTZ | TUM | LMU

TUM JOACHIM  
HERZ STIFTUNG



## Acknowledgements

theislab@  
Helmholtz Munich Computational Health Center

@giovp 

@g\_palla1 