nf-core   GHGA THE GERMAN HUMAN GENOME-PHENOME ARCHIVE   dkfz.

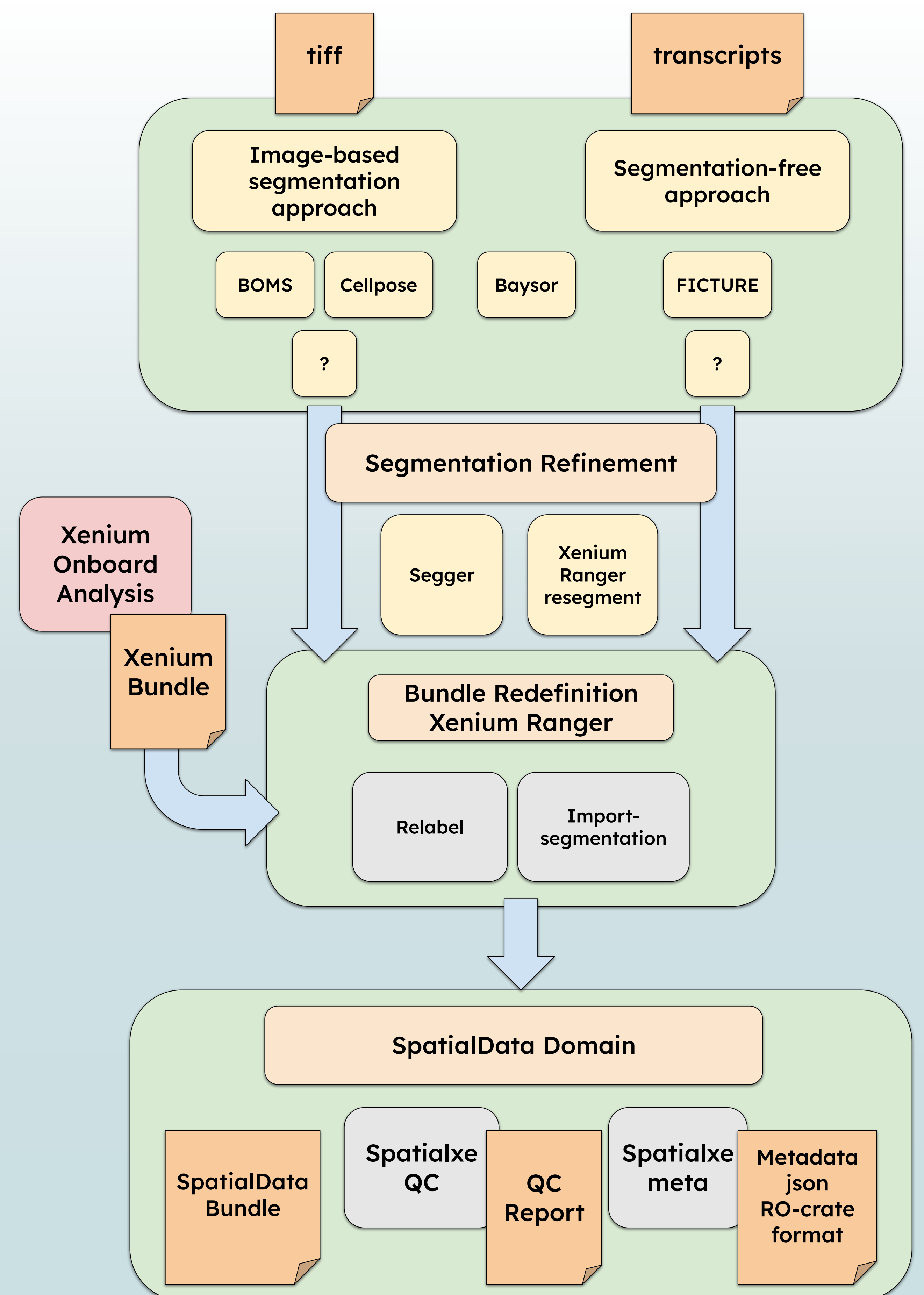# Standard processing and analysis of spatial Xenium in situ data with the nf-core workflow Spatialxe

Sameesh Kher [1,2], Florian Heyl [1,2], Oliver Stegle [1,2]

1 Division of Computational Genomics and Systems Genetics, Deutsches Krebsforschungszentrum (DKFZ), Heidelberg, Germany

2 German Human Genome-Phenome Archive (GHGA, W620), Deutsches Krebsforschungszentrum (DKFZ), Heidelberg, Germany

Spatial omics technologies represent a transformative approach in biological research, enabling the comprehensive analysis of molecular profiles within their native spatial context. By preserving the spatial relationships between cells, spatial omics technologies, such as 10x Xenium, offer critical insights into tissue architecture, cellular heterogeneity, and the microenvironment's role in a healthy and disease context. As there is an increase in demand for understanding spatial patterns to study diseases, there is a need for standardized and reproducible workflows. The nf-core community thus presents Spatialxe, a pipeline for the analysis of Xenium data.

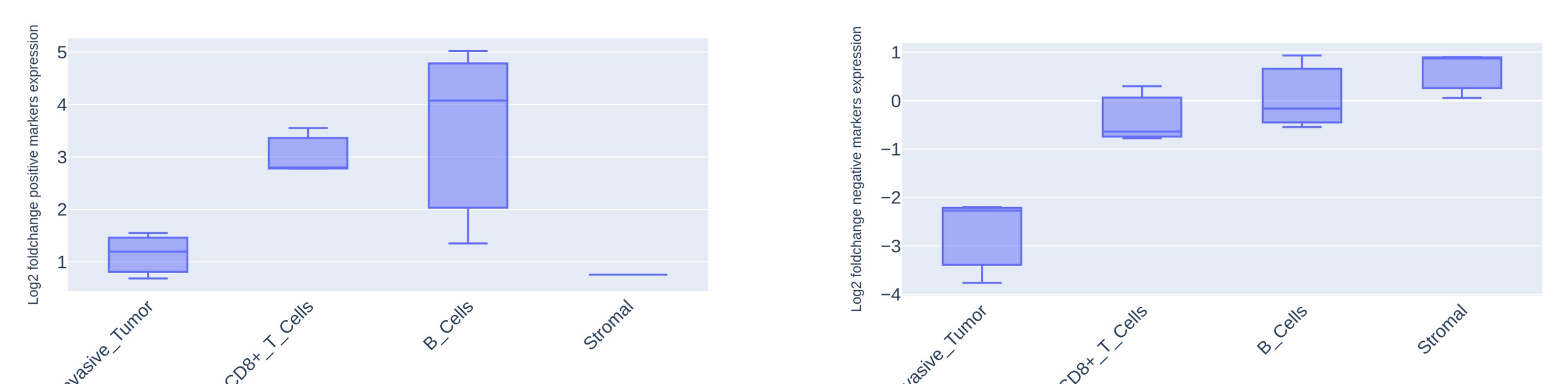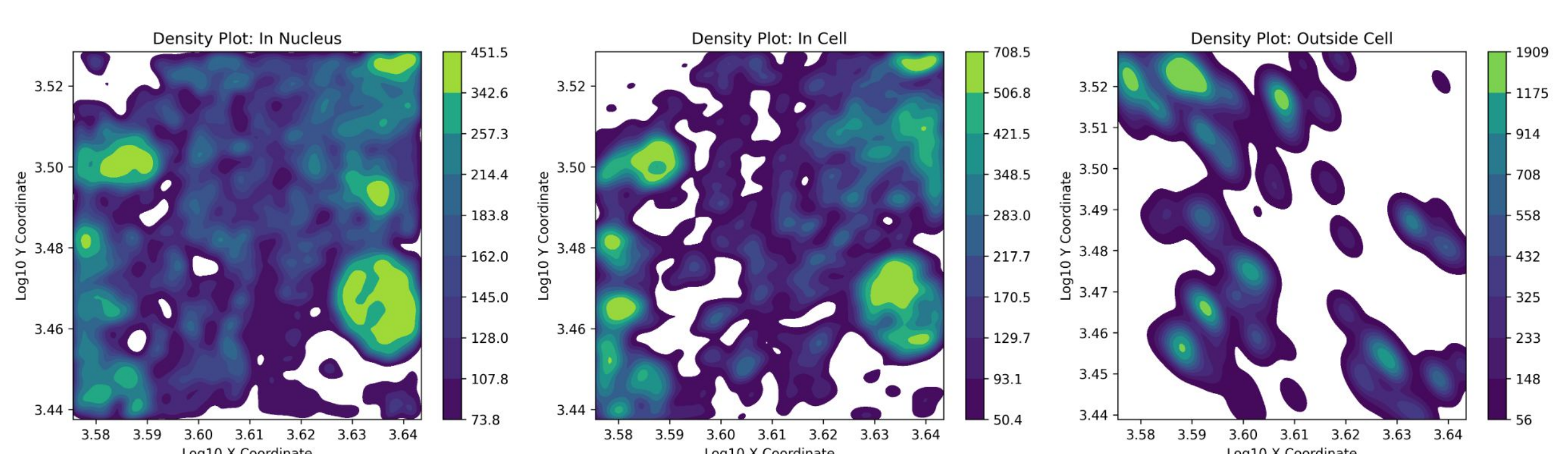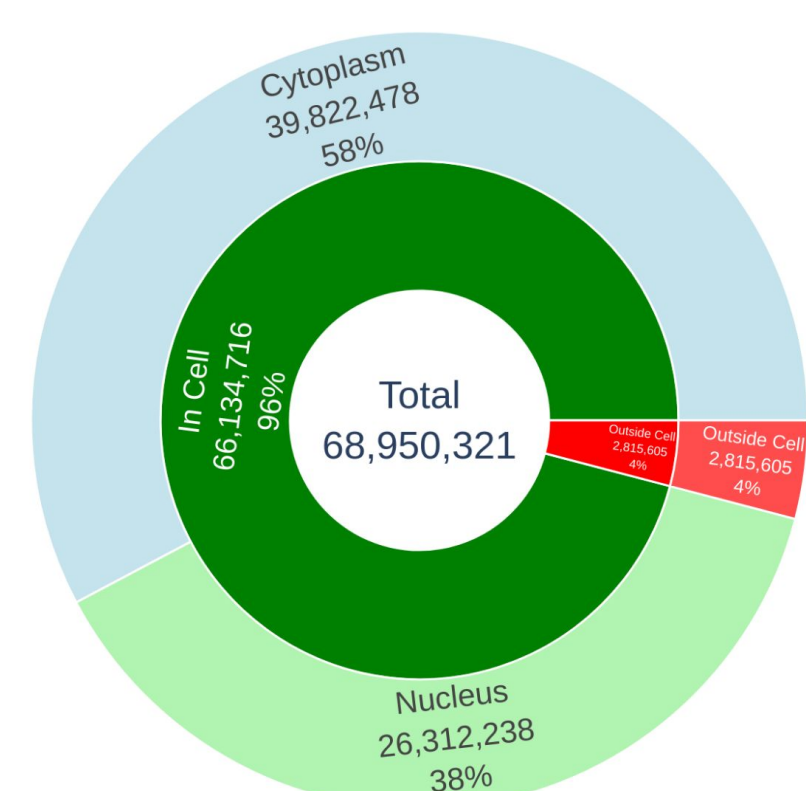- Spatialxe would support featured tools, such as Xenium Ranger.
- It generates a spatialdata object that can be used for further downstream analysis.
- We implemented a range of segmentation algorithms such as Cellpose, and Baysor.
- Spatialxe will be an extensive pipeline to cover not only standard processing but also single cell and spatial omics quality control and the generation of provenance metadata.
- The workflow will be deployed within the German Human Genome-Phenome Archive (GHGA) as the default analysis workflow for incoming Xenium data. The pipeline will be used to process and analyse Xenium data from primary tumor and metastases samples of hard-to-treat entities of colorectal cancer.

Spatialxe-QC comes with a user-optimized QC report, covering aspects such as:

- Transcript landscape
- Domain structure
- Expression data quality
- Marker assessment

- Domain inspection without ground truth allows for a first evaluation of the slide structure.
- Marker assessment should coincide with expectations and could signal a potential issue with marker selection.

Domain thickness: 0.7646