

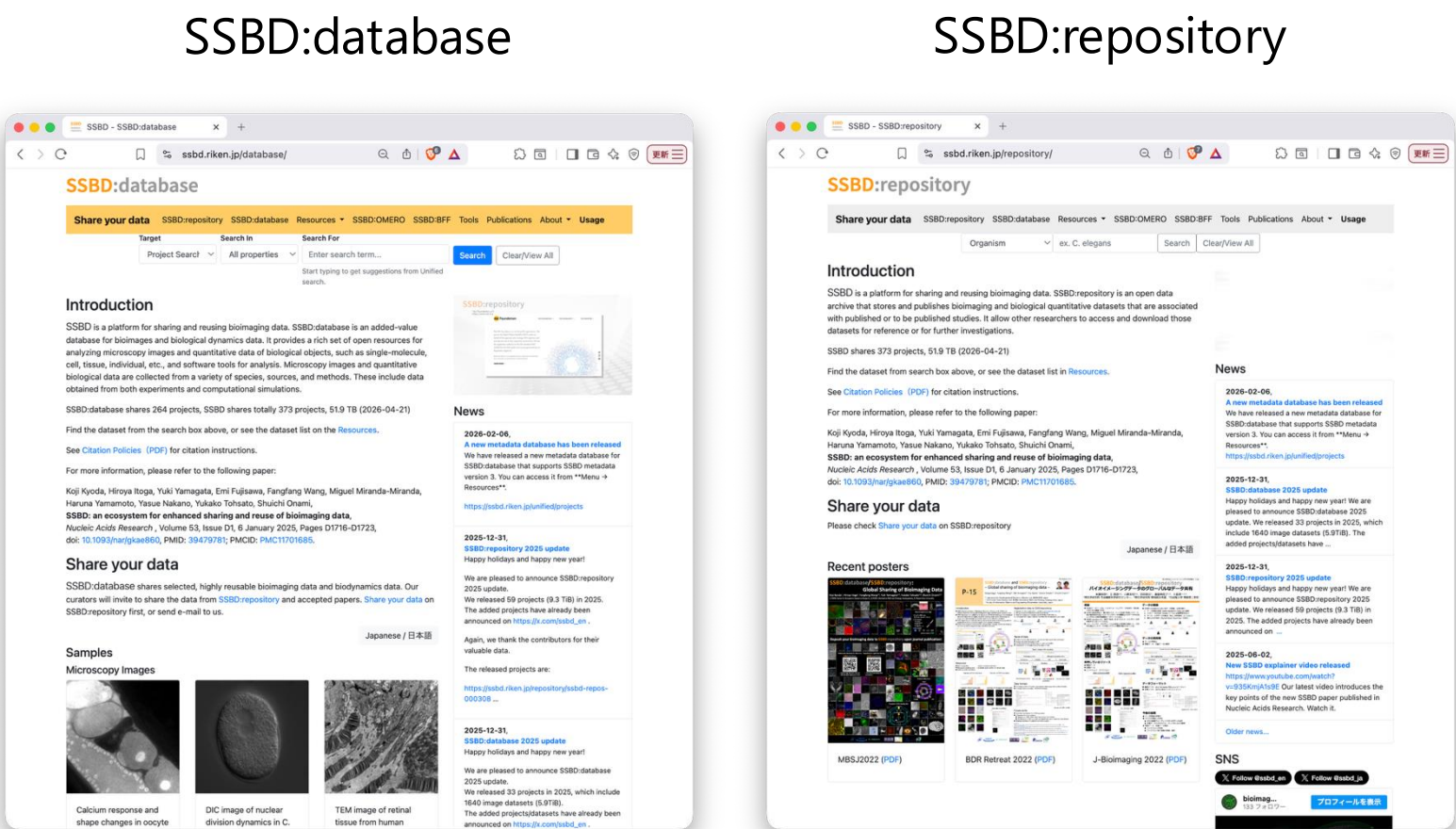
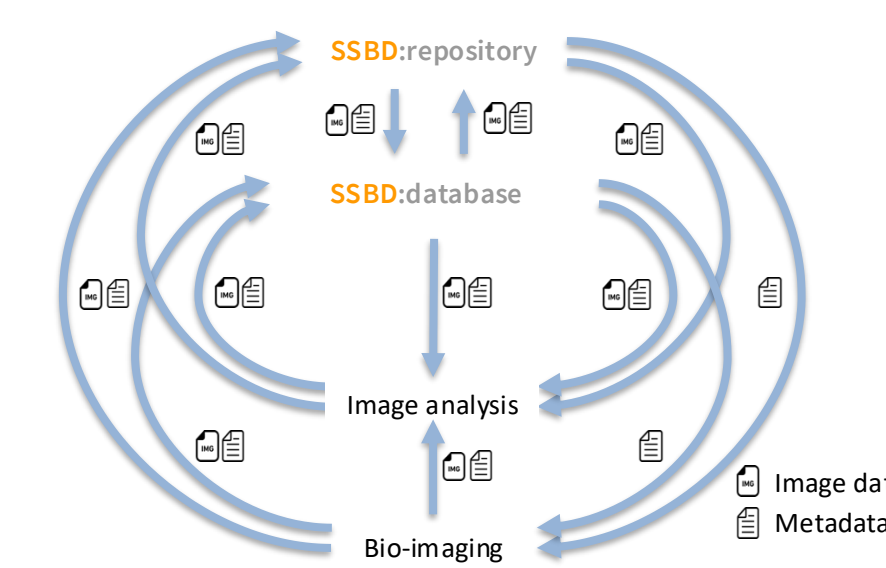
Building an International Bioimaging Data-Sharing Infrastructure with OME-Zarr and RO-Crate, and Extending It to Spatial Omics Data

Koji Kyoda¹, Hiroya Itoga¹, Yuki Yamagata^{1,2}, Emi Fujisawa¹, Yuzheng Wu¹, Tomohiro Nakamura¹, Haruna Yamamoto¹, Ko Sugawara¹, Shuichi Onami^{1,3} (¹RIKEN BDR, ²RIKEN BRC, ³RIKEN R-IH)

SSBD

- SSBD is a public repository for bioimaging data, developed since 2013.
- It consists of a two-tier system: data archive/repository and added-value database.
- SSBD serves as a core node in the GIDE (Global Image Data Ecosystem).

A bioimaging data sharing ecosystem

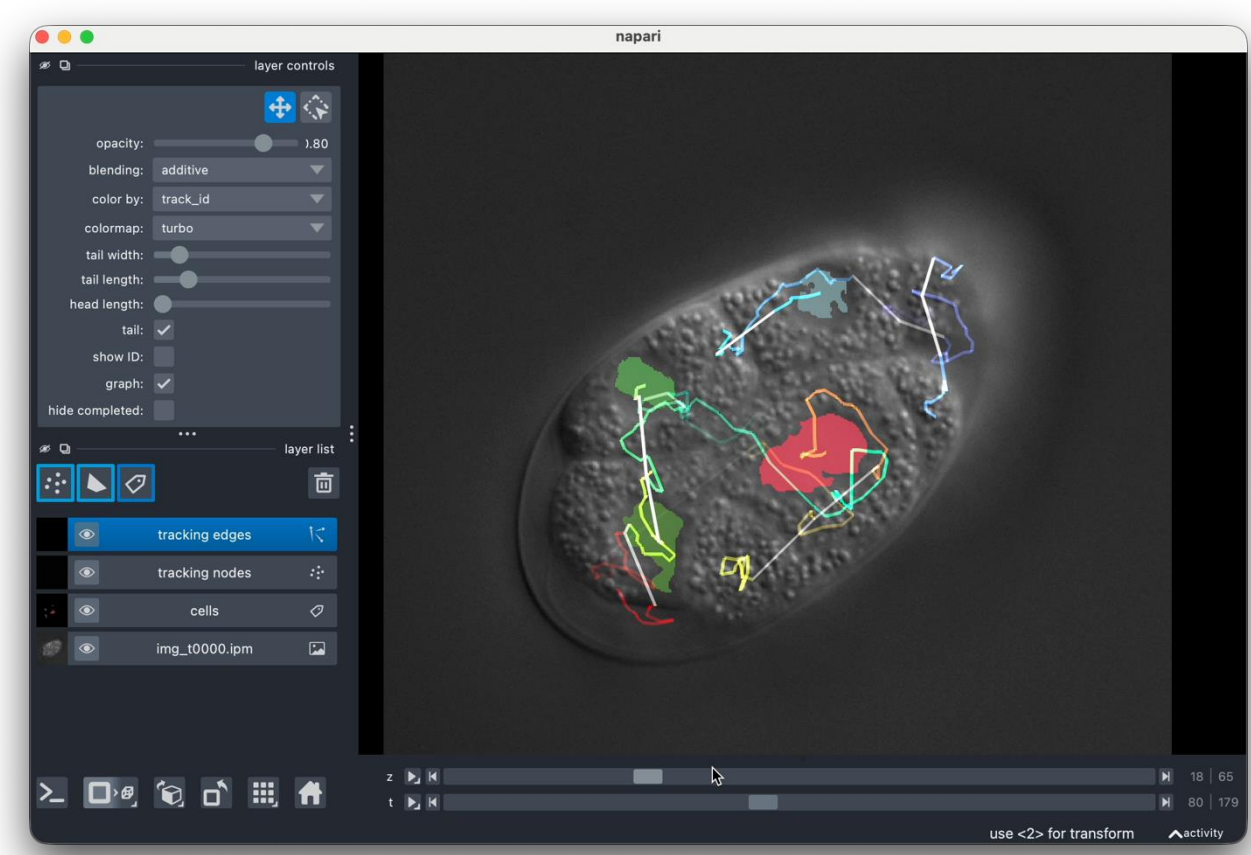


Advanced Data Sharing in OME-Zarr and GEFF

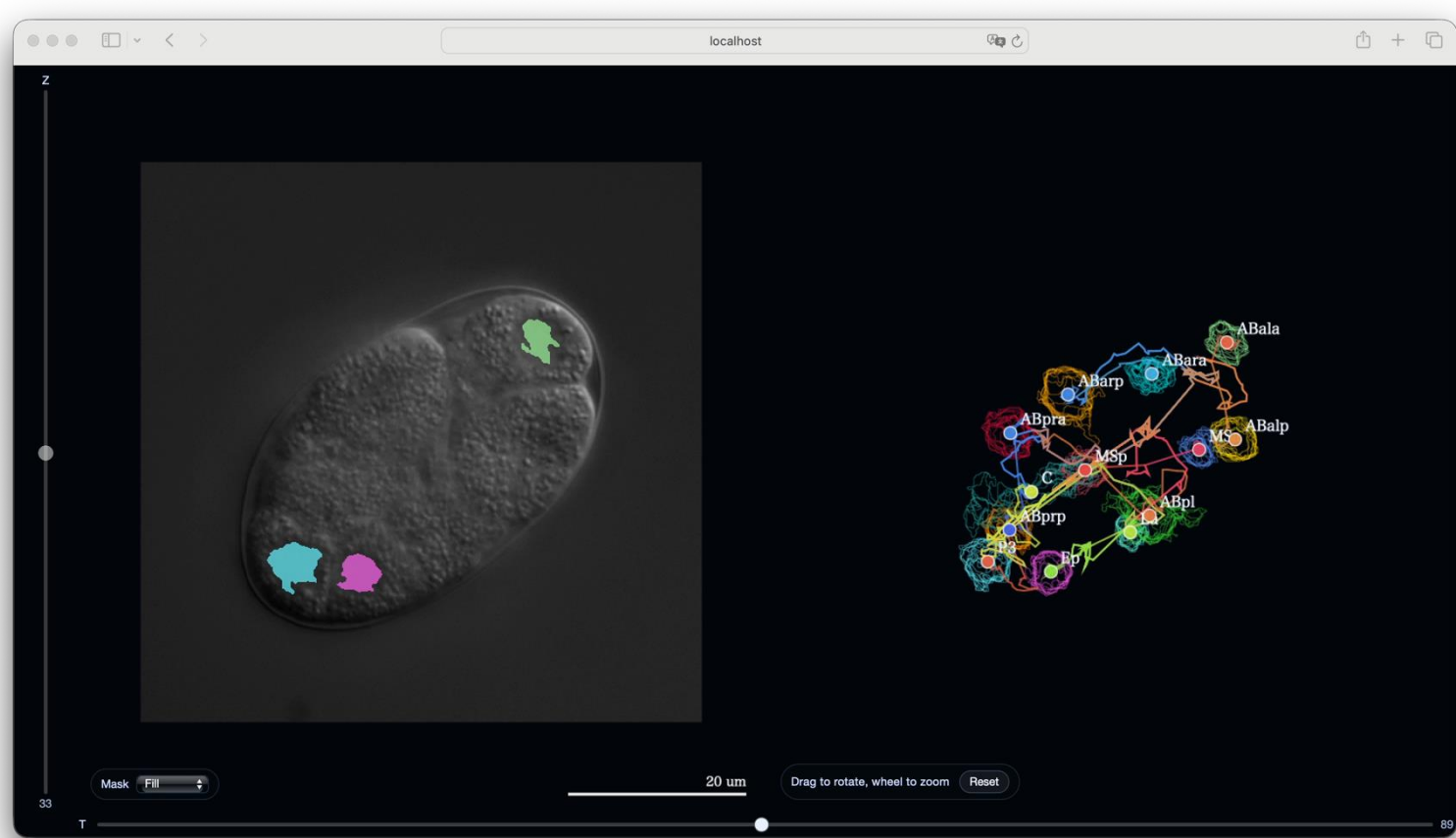
Moore et al., (2023)
Schwartz et al. (2026)

- Image and segmentation data are shared in OME-Zarr format.
- Tracking data are exchanged using Graph Exchange File Format (GEFF).
- A web viewer enables interactive visualization in the browser

Visualization in napari



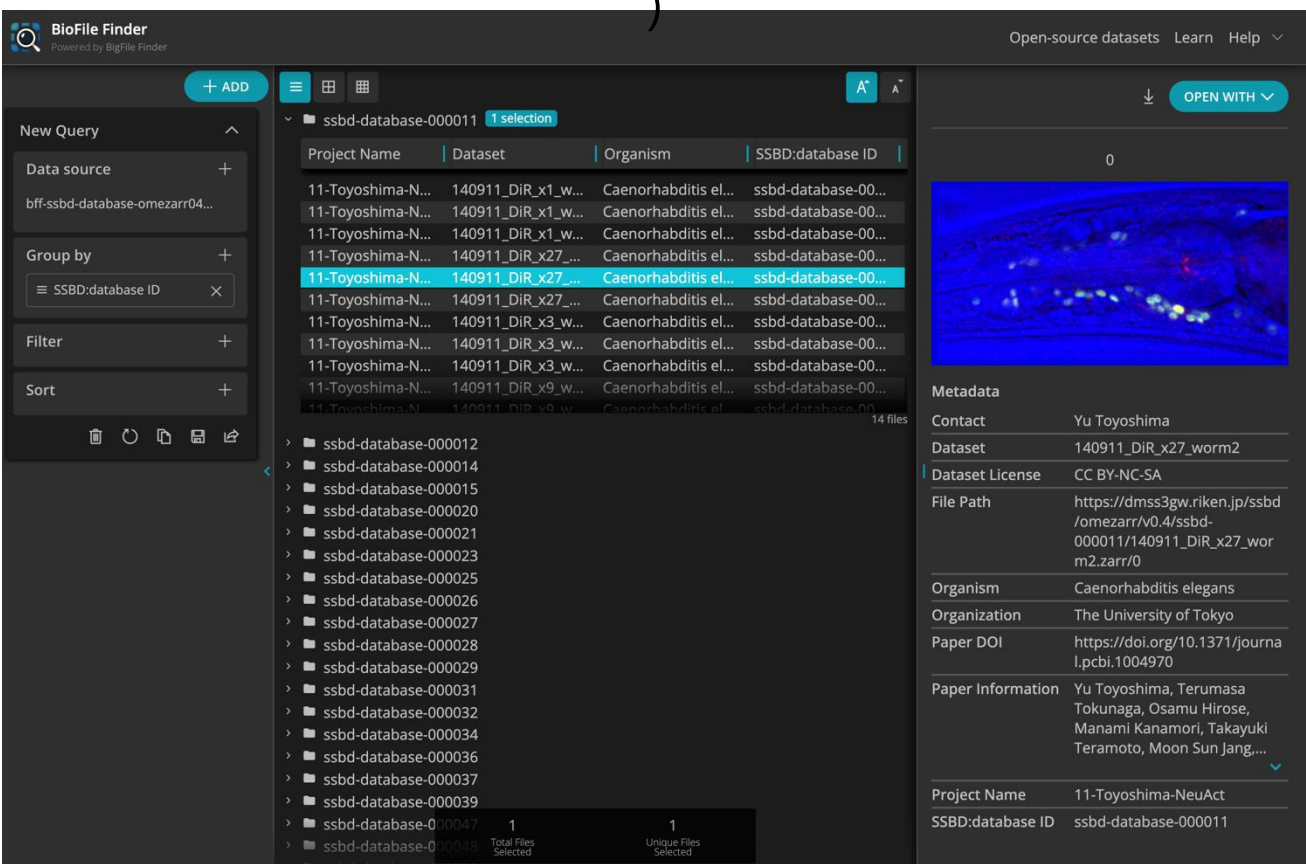
Visualization in web application



Gallery-Like Displays of OME-Zarr Data Collection

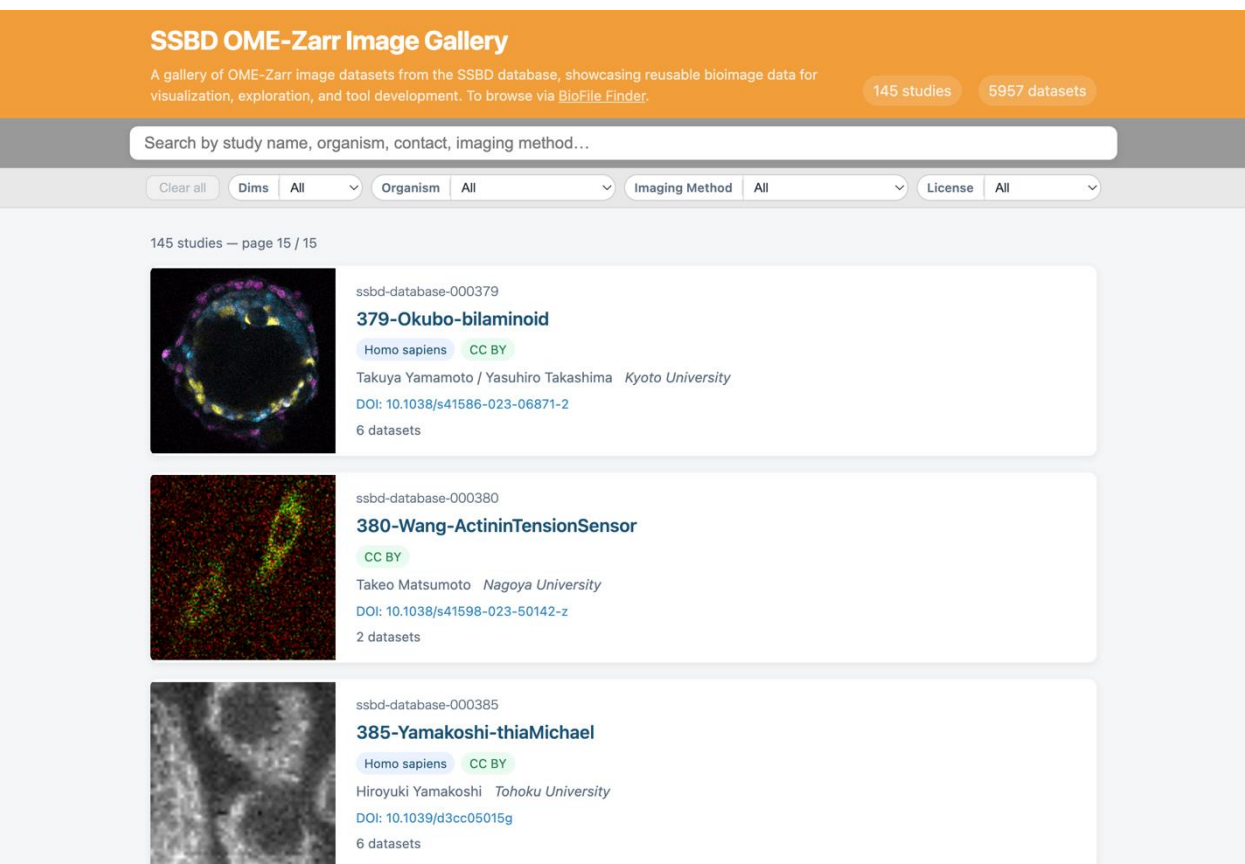
- BioFile Finder enables metadata-based discovery and access to distributed OME-Zarr datasets.
- Zallery organizes datasets into study-level galleries for intuitive browsing.
- Both tools improve data visibility, exploration, and reuse.

BioFile Finder (<https://ssbd.riken.jp/ssbd-bff/>)



Maharry et al. (2026)

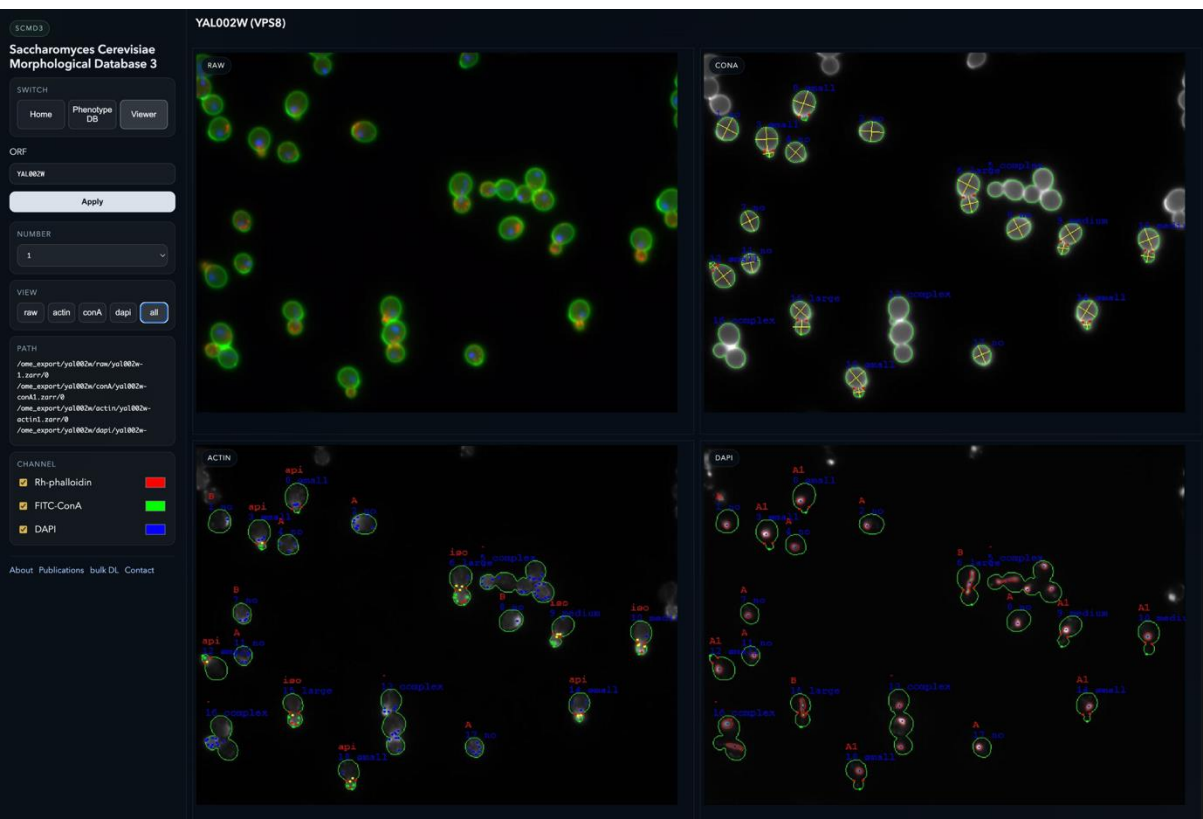
Zallery (<https://openssbd.github.io/zallery/>)



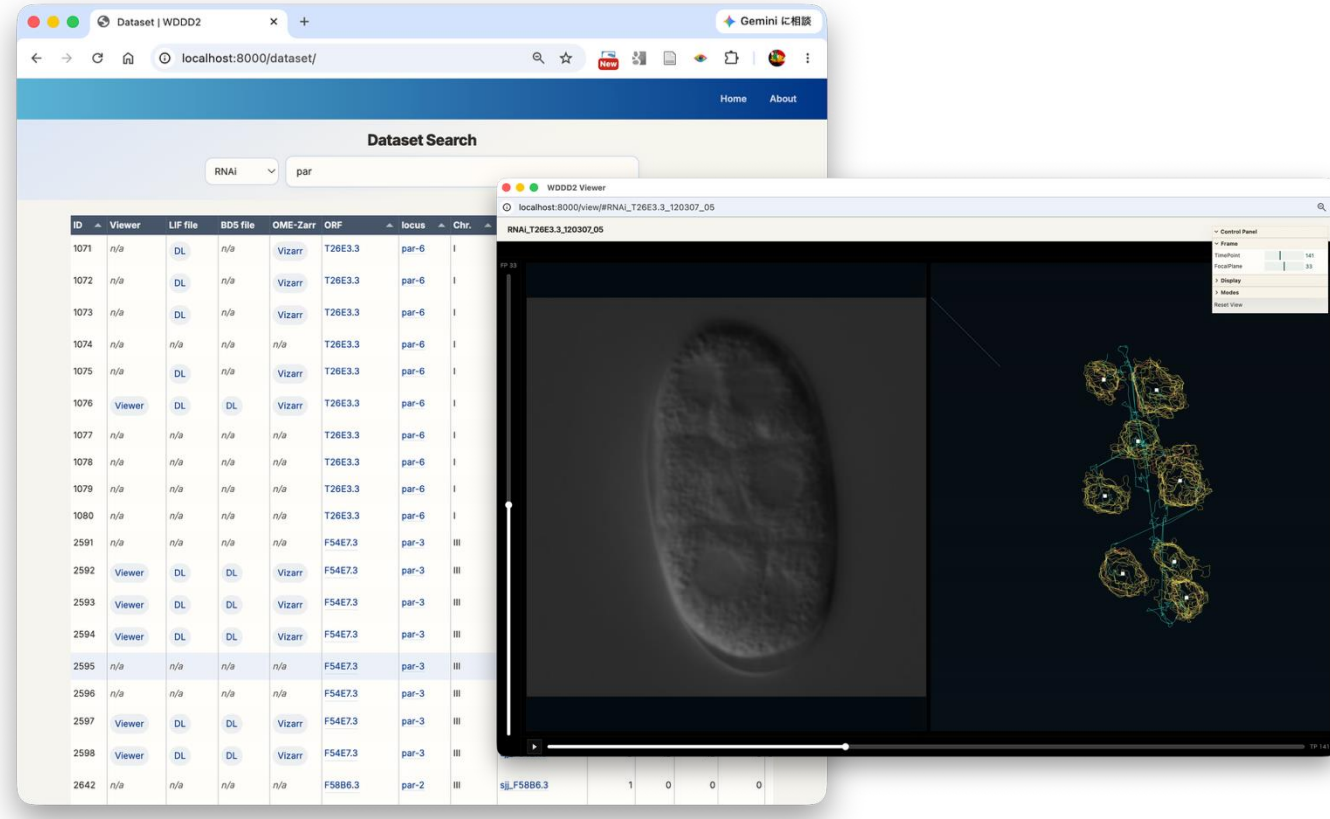
Domain-Specific Databases Based on OME-Zarr

- SCMD3: 2D morphological images and phenotype data in *S. cerevisiae*
 - haploid deletion mutants for non-essential genes (4,718)
 - diploid heterozygous deletion mutants for essential genes (1,112)
- WDDD2: 4D microscopy images and division dynamics in early *C. elegans* embryos
 - RNAi-treated embryos for all essential embryonic genes (263)

SCMD3



WDDD2



GIDE Data Search with RO-Crate

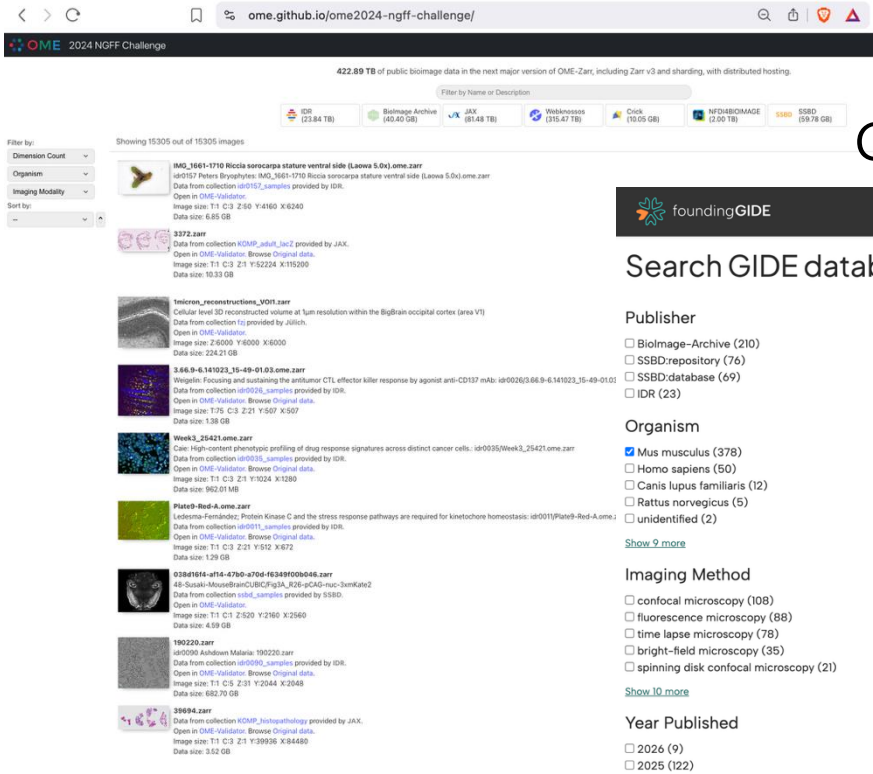
- RO-Crate provides a common, machine-readable metadata framework across repositories
- Harmonized metadata enables cross-repository search (BIA, IDR, and SSBD).
- Linked metadata improves discovery, interoperability, and reuse of bioimaging data.

```
{
  "id": "https://idr.openmicroscopy.org/study/idr0167/",
  "entry": {
    "type": "Dataset",
    "name": "Predicting cell cycle stage from 3D single-cell nuclear-stained images",
    "author": {
      "id": "https://orcid.org/0000-0002-8501-8028",
      "type": "Person",
      "name": "Eva Nichols",
      "givenName": "Eva",
      "familyName": "Nichols"
    },
    "description": "The cell cycle governs proliferation of all eukaryotic cells [...]",
    "datePublished": "2026-03-30",
    "license": "https://creativecommons.org/licenses/by/4.0/",
    "publisher": {
      "id": "https://idr.openmicroscopy.org/",
      "type": "Organization",
      "name": "Image Data Resource"
    },
    "about": {
      "type": "BioSample",
      "name": "cell",
      "taxonomicRange": [
        {
          "id": "http://purl.obolibrary.org/obo/NCBITaxon_10890",
          "type": "Taxon",
          "scientificName": "Mus musculus"
        }
      ],
      "measurementMethod": [...]
    }
  }
}
```

Hartley & Sherwood., (2026)

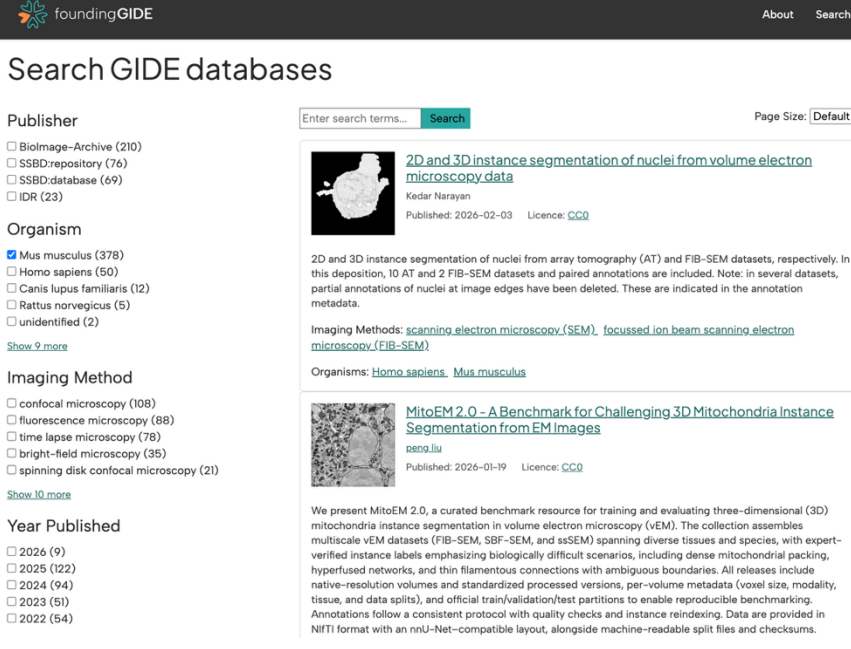


OME 2024 NGFF Challenge



Moore et al., (2025)

Gide Search Portal

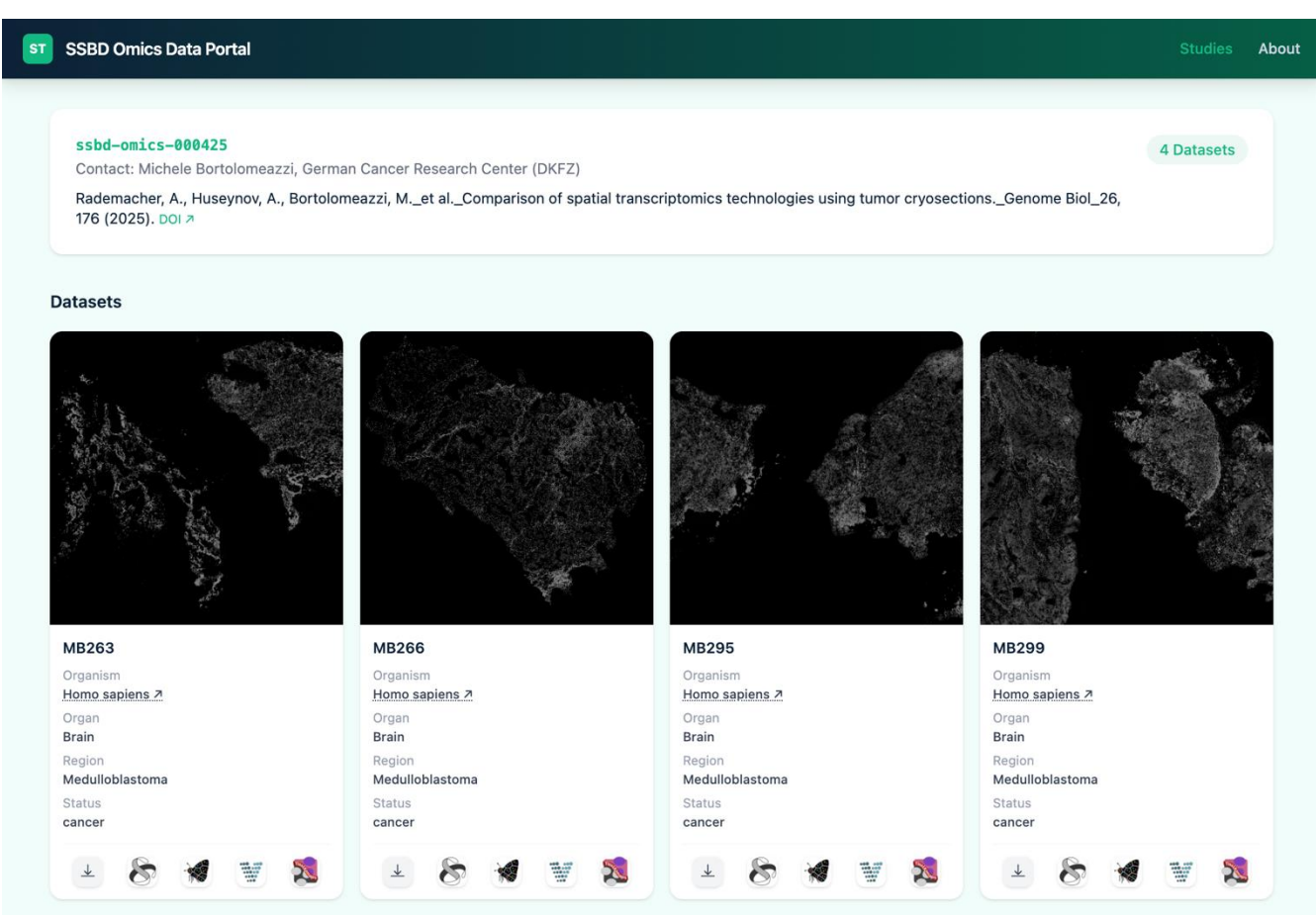


<https://www.gide-project.org/portal>

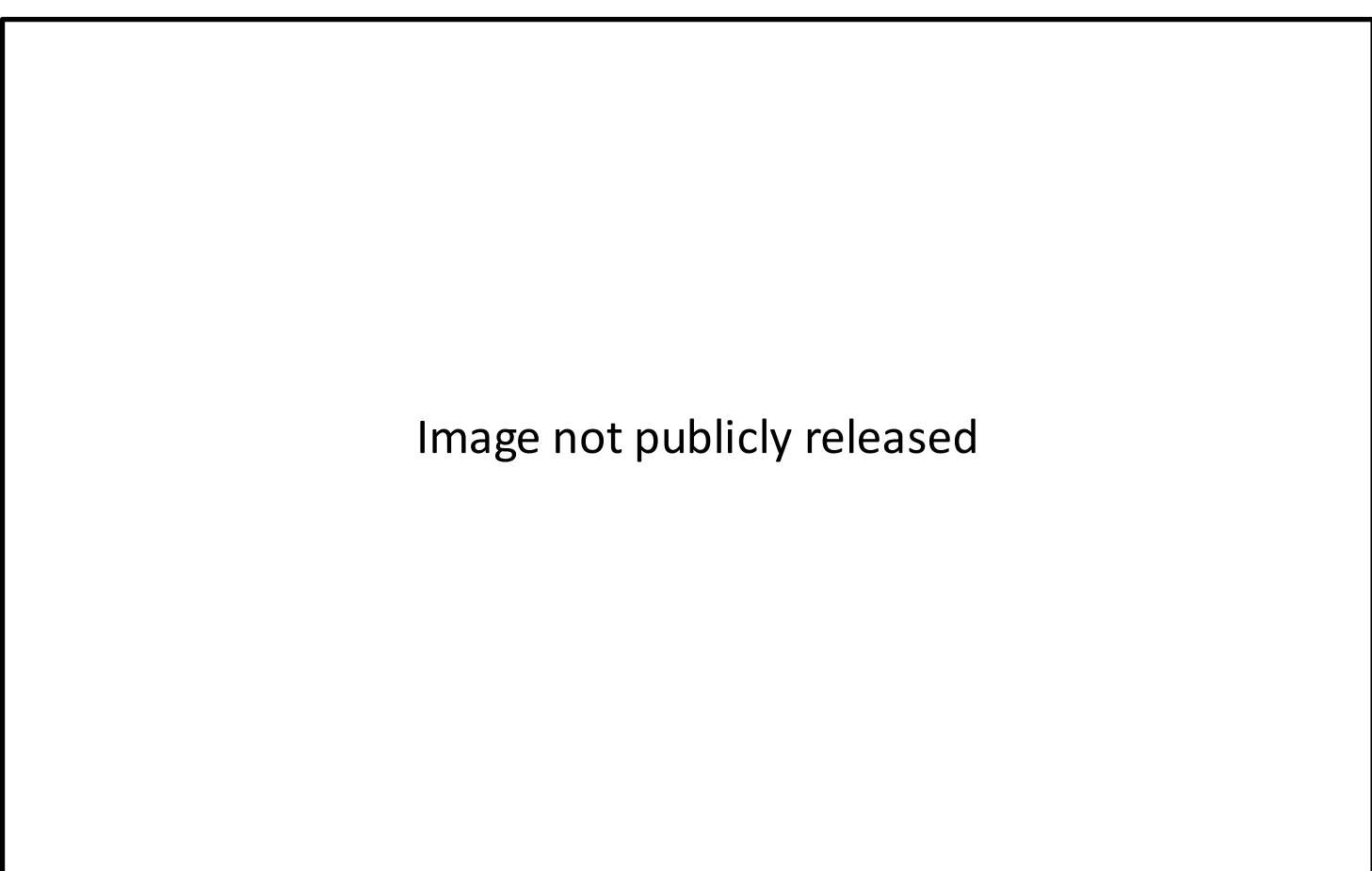
Spatial Omics Data Portal

- We are developing a public repository for spatial omics data.
- The portal will support both imaging- and sequencing-based technologies.
- Interactive visualization tools are also being developed for data exploration.

SSBD Omics Data Portal (tentative)

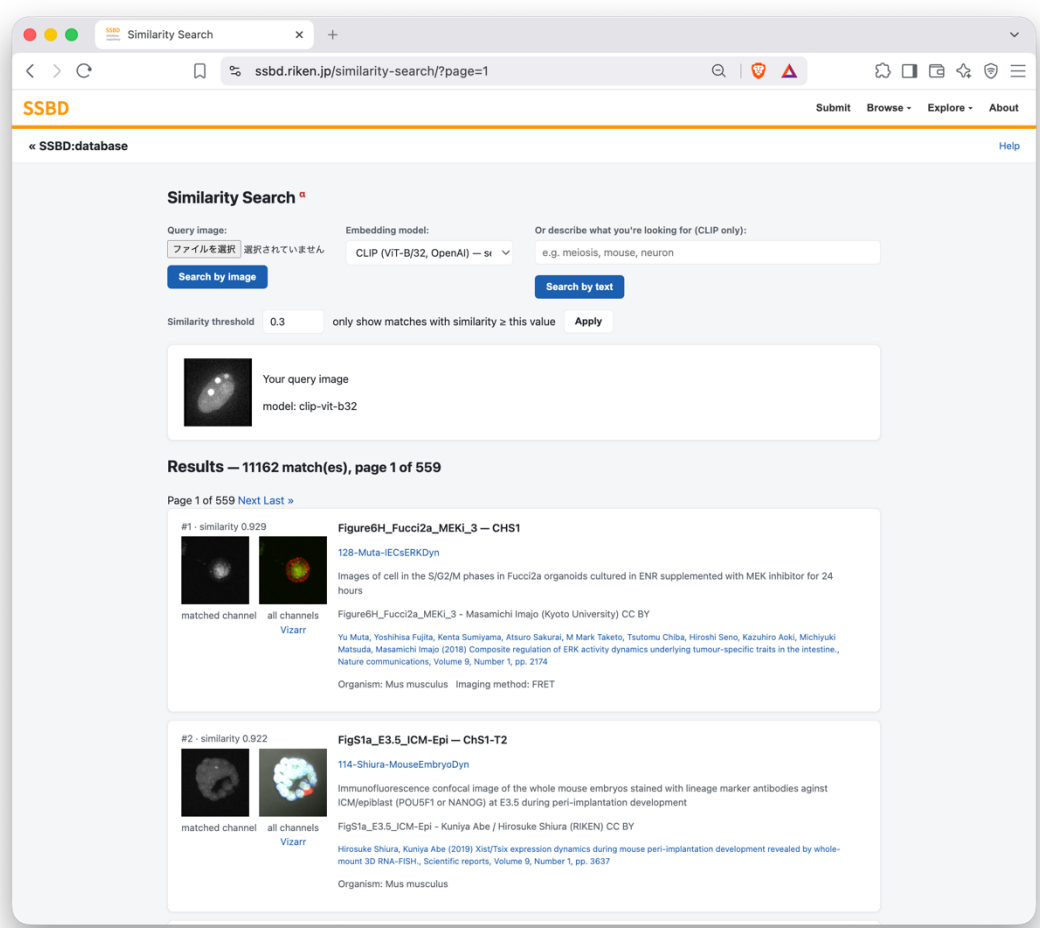
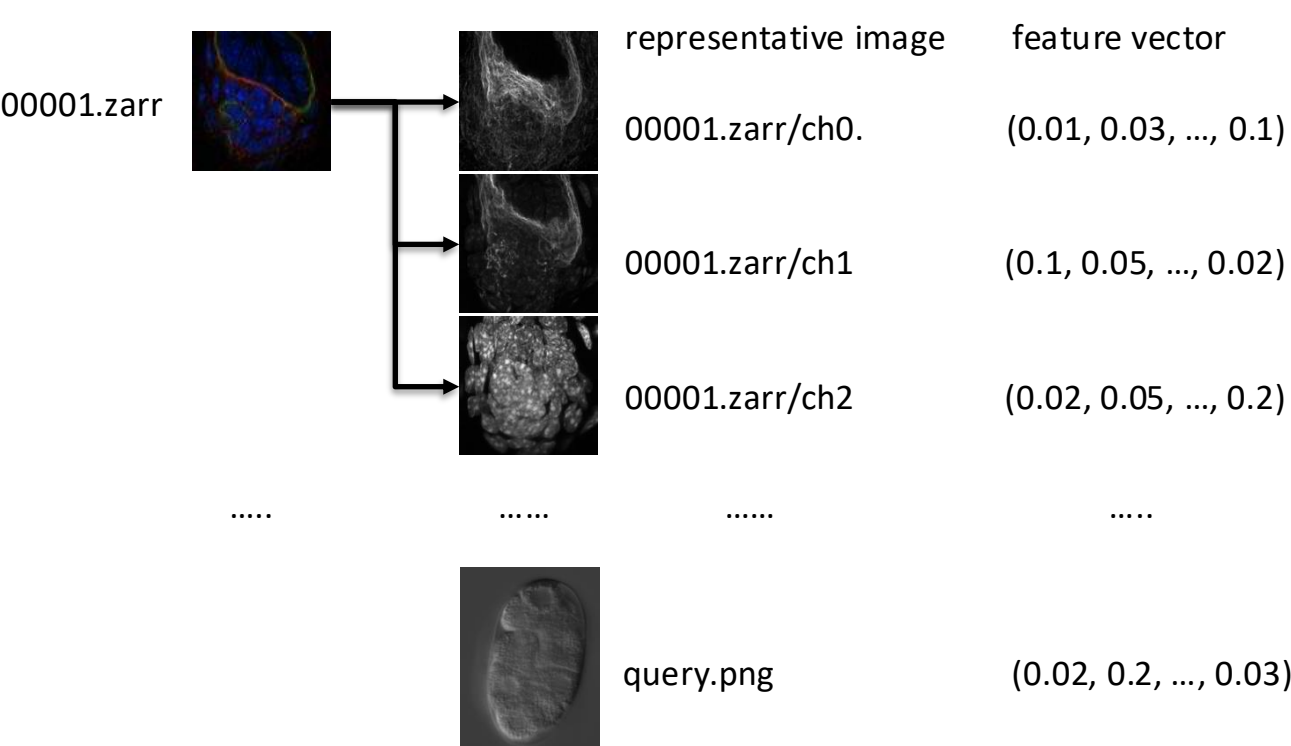


SOVIA: Spatial Omics Viewer for Interactive Analysis



Similarity Search: Query-by-Image Discovery

- Search visually similar bioimages in SSBD:database using CLIP, DINOv2, and uniDINO feature embedding.
- The current index covers 5,954 OME-Zarr image sets and 11,162 image channels.
- We are exploring its use of dataset discovery, cross-repository search, and metadata-assisted annotation.



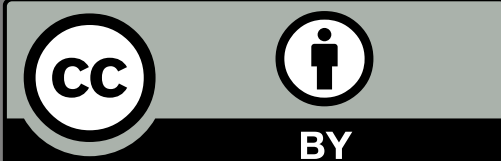
<https://ssbd.riken.jp/similarity-search/>

Summary

- SSBD is making bioimaging data easier to find, access, and reuse.
- OME-Zarr, BFF, and Zallery support scalable data exploration.
- RO-Crate supports harmonized metadata and cross-repository search via GIDE.
- SSBD is expanding into spatial omics with portals, viewers, and conversion tools.
- Query-by-image search for discovering similar bioimaging datasets.

References

- Kyoda et al. (2025). *Nucleic Acids Res.* **53**, D1716-D1723.
- Moore et al. (2023). *Histochem. Cell Biol.* **160**, 223-251.
- Schwartz et al. (2026) *J. Open Source Softw.* **11**, 10143.
- Maharry et al. (2026) *Nat. Methods* **23**, 1256-1258.
- Onami et al. (2025). Zenodo [<https://doi.org/10.5281/zenodo.16794787>]
- Moore et al. (2025). Zenodo [<https://doi.org/10.5281/zenodo.16735915>]



Licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) ©2026 Koji Kyoda (RIKEN BDR)

