

GBI 2022 IDR Workshop

The presentation and a PDF version of the workshop are available at <https://downloads.openmicroscopy.org/presentations/2022/GBI>

Software versions used for this workshop:

- OMERO: 5.6.3
- OMERO.web: 5.12.0
- OMERO.iviewer: 0.11.1
- OMERO.mapr 0.4.1
- omero-guides: 2021.11.09
- Bio-Formats: 6.5.1
- Fiji/ImageJ: 2.0.0-rc-69/1.52p

Summary

Introduction to IDR

Searching in IDR

1. From publication to IDR study, viewing images and ROIs
2. Genes search, cross-linking to other Databases, metadata ecosystem
3. Compounds search, effective doses metadata exploration

Search images in IDR based on metadata via API and notebooks (demo and slides)

1. Dose-response curves for compounds screened in SARS-CoV-2 study.
2. Compounds investigation - overview of chosen compounds usage with image examples.
3. Cross-study search using an externally mined information in a notebook

Walkthrough

- **Searching in IDR:**
 1. **From publication to IDR study, viewing images and ROIs**
 - Start from PLOS Biology or [PMC](#). Search for two terms, “Blin” and “Lowell”. First result will be a paper from Blin et al from Sally Lowell’s lab published in PLOS Biology in 2019.
 - Navigate to the Data availability section and then go from there to the starting page at <http://idr.openmicroscopy.org/> and search for a “Name (IDR number)”. (the default option). Enter the “62” into the second search box. Accept the suggestion and click on the thumbnail. This gives you all the images contained in the idr0062 study..This workflow shows one of the main uses of IDR for data visibility, Quality Control for publishers and community.

- In the [idr0062 study](#), on the image <http://idr.openmicroscopy.org/webclient/?show=image-6001239> we can see the attached tiff file with ROIs. These ROIs were extracted into OMERO ROIs (masks) which can be seen in OMERO.iviewer (by double-clicking on the thumbnail of the image in the central pane).
- Click on the [Experiment](#) (named idr0062-blin-nuclearsegmentation/experimentA) and expand the “Attributes” harmonica in the right-hand pane. Find and click on the DOI of the paper link.

2. Genes search, cross-linking to other Databases, metadata ecosystem

- Start on <http://idr.openmicroscopy.org/> searching for “Gene” with a value “Baz1a”. This is Bromodomain adjacent to zinc finger domain protein 1A” and is a component of some ISWI chromatin remodeling complexes
- Search results are not case sensitive, so that ensures different gene naming conventions don’t affect the search results. In this case, that’s demonstrated by search results for both human and mouse genes.
- Navigate to idr0097 in results (last item) and click on “More” to see the full view and metadata of the images.

Search by:

Found **194** images with **Gene: Baz1a/BAZ1A** in **8** studies

Baz1a

Study ID	Organism	Image count	Title	Sample Images	Link
idr0018A	Mus musculus	146	Histopathology of mouse knockouts		more...

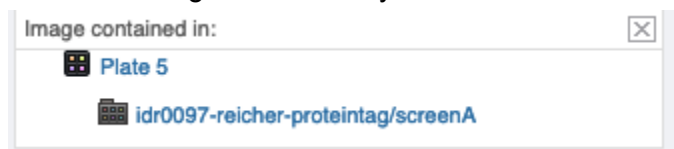
BAZ1A

Study ID	Organism	Image count	Title	Sample Images	Link
idr0009A	Homo sapiens	14	Genome-wide RNAi screening identifies hu...		more...
idr0010A	Homo sapiens	2	RNF168 binds and amplifies ubiquitin con...		more...
idr0012A	Homo sapiens	2	Clustering phenotype populations by geno...		more...
idr0013A	Homo sapiens	8	Phenotypic profiling of the human genome...		more...
idr0013B	Homo sapiens	8	Phenotypic profiling of the human genome...		more...
idr0093A	Homo sapiens	9	High content genome-wide siRNA screen to...		more...
idr0097A	Homo sapiens	5	Pooled protein tagging, cellular imaging...		more...

- This will bring you to the OMERO.web user interface, showing a tree with search results. These are studies that include images related to the Baz1A genes.
- Only the images associated with Baz1a are shown, but there are more images inside each study.
- Navigate to the top Screen and find in the right-hand pane under “Attributes” harmonica the link to the publication. This is a study of protein localization published by Reicher et al in Genome Research in 2020.
- Select one single image inside that screen. Again, expand the “Attributes” harmonica. This contains the most important metadata associated with this particular image. We have noted that authors found a phenotype when the authors disrupted the gene and we also have the intron-targeting guide RNA used by the authors.
- Navigate to the whole study via click on link “Show parent



Project&Dataset” – **Show parent Projects & Datasets** in the right-hand pane and then click on the Screen (named idr0097...). In this way you will get a view of all images in the study.




- This is a display that represents the original screening data where we record information of the wells, number of images per well,
- Open Plate 5, click on G1 and then see the images inside that “G1” well at the bottom of the central pane.
- Click back on the well G1 in the grid and open Tables accordion to reveal further metadata - these are all the metadata associated with the well G1.

3. Compounds search, effective doses metadata exploration.

- Still in OMERO.web, click on the top link “Compounds” to get to a search environment. Note that this does not give you a full list of compounds hosted in IDR, instead, you have to start typing a compound name and use the autocomplete function for a help. Start typing “Remdesivir”. Into the dialog, select the suggestion. Expand the tree with search results (over 200 images found). This is a high content screen from Phil Gibbon’s group published in Scientific Data that was the first published drug screen trying to identify compounds that inhibited SARS-CoV-2 infection in human cells
- Go to first plate first image and expand the Attributes harmonica in the right-hand pane. There, find the Compound block and click on Remdesivir

icon bringing you to Pubchem.

Cell Lines	
Added by: Public data	
Cell Line	Caco-2

Compound	
Added by: Public data	
Compound Name	Remdesivir 

Compound supplementary	
Added by: Public data	
Identifier	CTRL_Remdesivir
Concentration (microMolar)	20

- This is a link to the compound summary in the <https://pubchem.ncbi.nlm.nih.gov/compound/121304016>.
- The DPC inhibition parameter is available too (under “Others” section of the “Attributes” harmonica. This is data submitted by the paper authors.

Organism	Severe acute respiratory syndrome coronavirus 2
----------	---

Others	
Added by: Public data	
Quality Control	pass
Percentage Inhibition (DPC)	96.81
Hit Compound (over 75% activity)	No
Cells - Number of Objects	25159
Phenotype Annotation Level	well level
Channels	cell body

Tables	
---------------	--

Show the correspondence between dark thumbnails with low DPC and bright thumbnails with high DPC.

4. **Search images in IDR based on metadata via API and notebooks (demo and slides)**
 - a. Dose-response curves for compounds screened in SARS-CoV-2 study.
 - i. [Study of Remdesivir compound case using a notebook](#)

1. Also linked from [idr0094 screen](#) - expand the Attributes harmonica and scroll to bottom to start a Binder build.
- ii. [Dose response curves of all compounds in Shiny app](#)
- iii. This example shows that the submitter should also submit, if possible, complex analytical data. The analytics will be exposed, when suitable, allowing viewers to partially reproduce some findings associated with the study or dynamically explore the analytical data. This also allows users e.g. image analysts wishing to re-analyse the data with different methods to compare their findings with the original analytical results.
- b. [Compounds investigation](#) - overview of chosen compounds usage with image examples.
- c. [Cross-study search](#) using an externally mined information in a notebook
 - i. Note also the last slide of the presentation

If you wish to explore the notebooks Ad b. and c. yourself, start [Binder with notebook](#)