



Image Data Resource (IDR) Jean-Marie Burel



University
of Dundee

ABiS-GBI course

05 July 2023

@openmicroscopy, @IDRNews

GLOBAL
BIOIMAGING
growing collaboration



EMBL-EBI



The Image Data Resource (IDR): a scalable resource for FAIR biological imaging data

- **Public access**
- **Reference datasets** - complete datasets containing molecular and functional annotations, associated with an existing or upcoming publication.
- **Study integration** - integrating studies or datasets with other datasets via **genes**, **compounds** or **phenotypes**.
- **Curated metadata**
- **Cloud re-analysis**

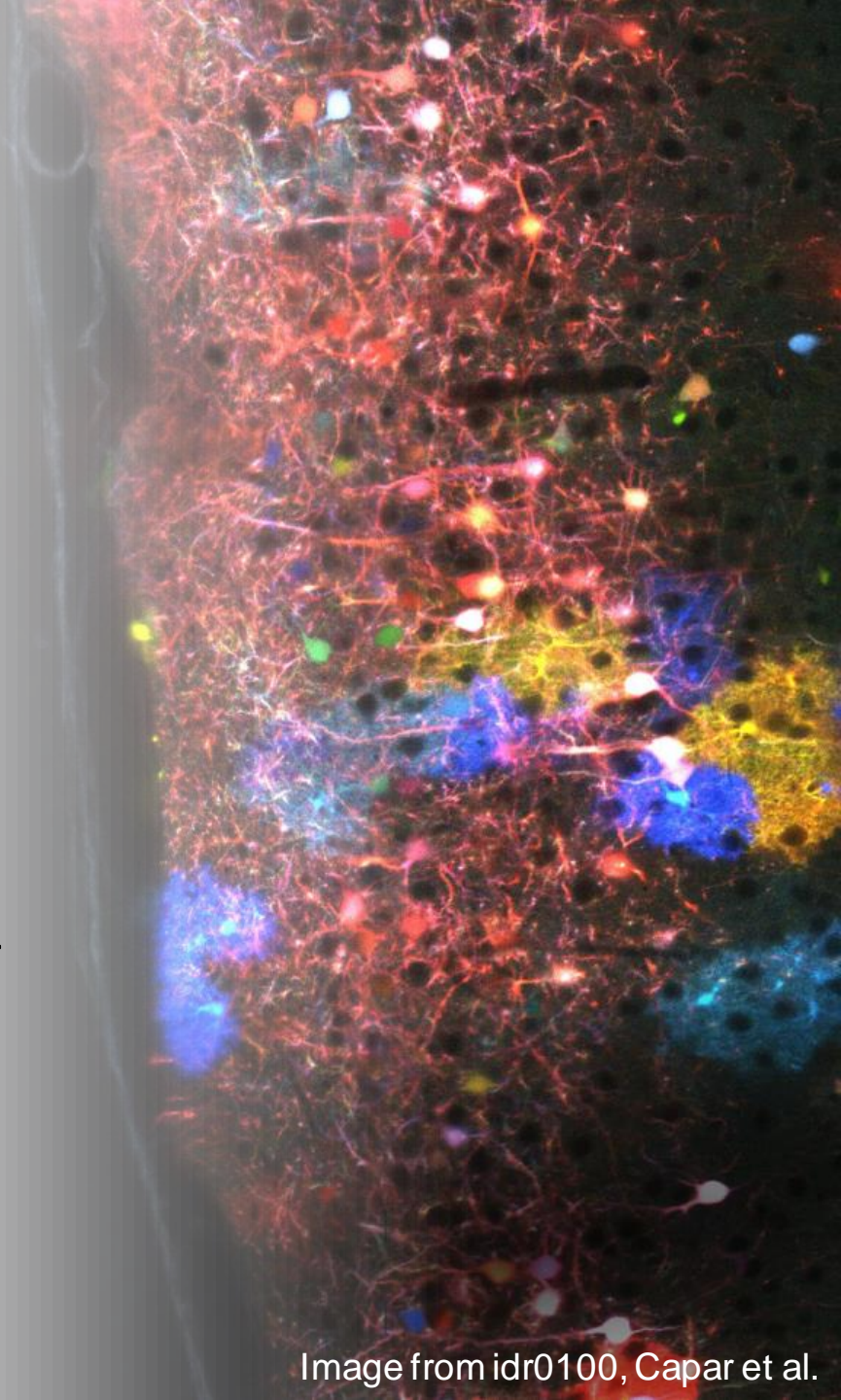
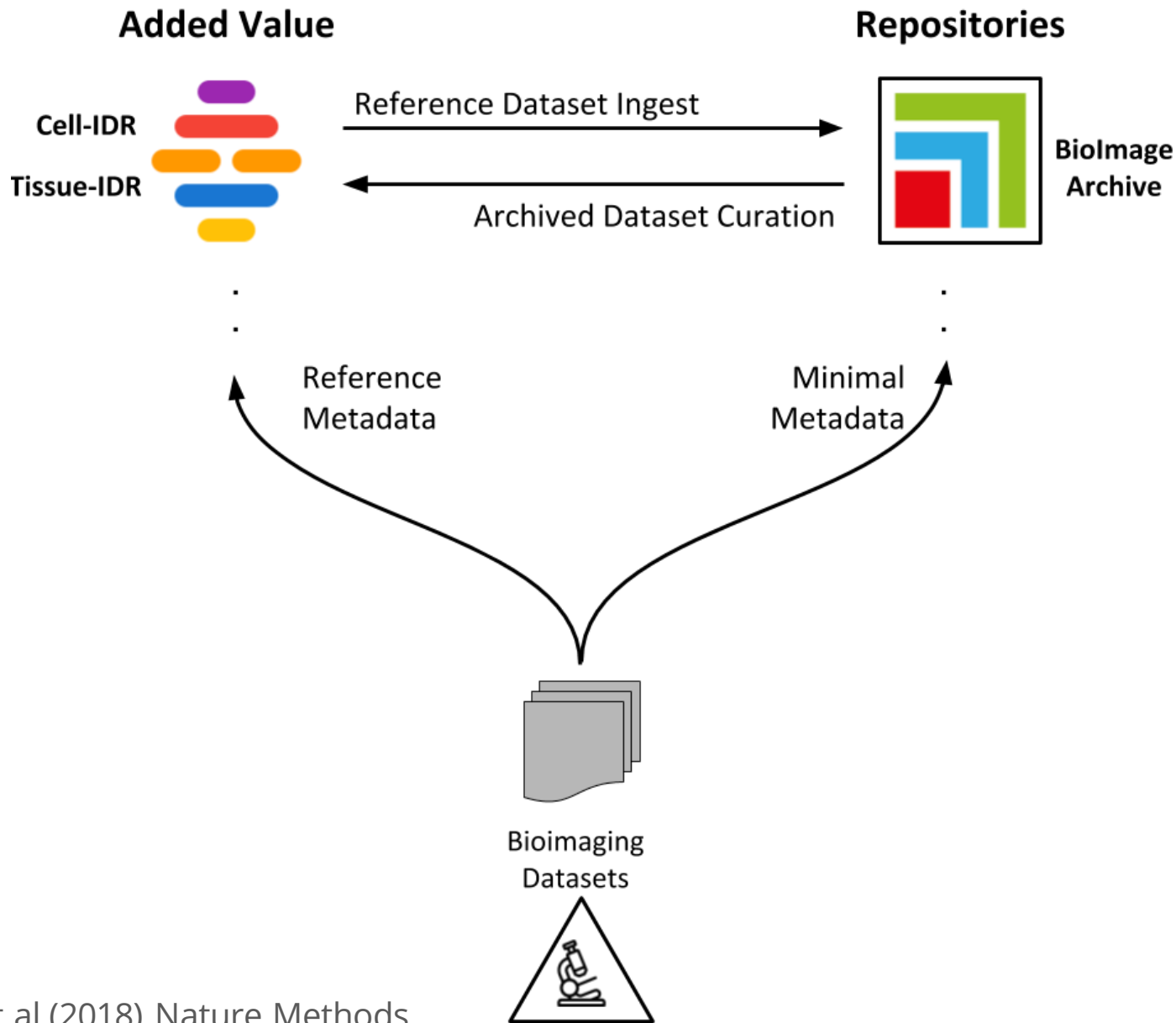




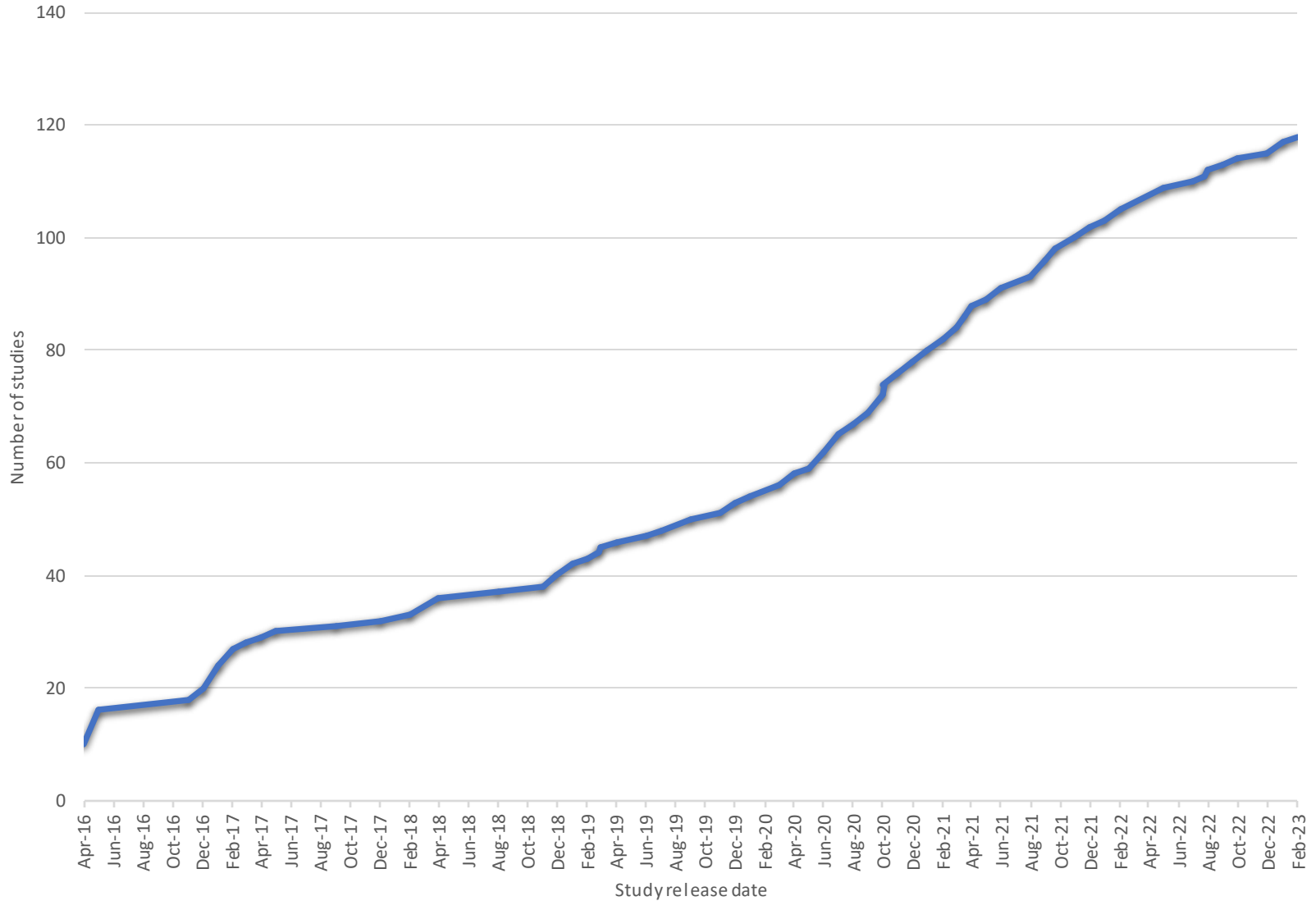
Image Data Resource (IDR):

- **is a public version of OMERO**
- **is READ-ONLY server**

Building the BioImage Ecosystem



Published studies in IDR



June 2023

Raw data size	361 TB
Image files	32 M
5D Images	13 M
2D Planes	112 M
Organisms	71
Genes	89 K
Antibodies	12 K
Compounds	40 K

>115 cross-published studies



**cross-referenced via
accessions and DOIs**

A high-content RNAi screen reveals multiple roles for long noncodin...

Data availability

A reporting summary for this article is available as a [Supplementary Information file](#). Sequencing data are available in the ArrayExpress database (<http://www.ebi.ac.uk/arrayexpress>) with the accession codes [E-MTAB-7432](#) (RNA-seq), [E-MTAB-7418](#) (CHART-seq) and [E-MTAB-7419](#) (CUT&RUN). The imaging data have been submitted to the Image Data Resource (<https://idr.openmicroscopy.org>) under IDR accession number [idr0056](#). The source data

IDR CELL - IDR TISSUE - IDR ABOUT SUBMISSIONS

Home

Attribute: Name (IDR number) Operator: contains Value: idr0056

Search **name** contains **idr0056** found **220950** images in **3** experiments/screens

ID	Images	Publication Title
idr0056B	115110	A long noncoding RNA regulates microtubule behaviour during mitosis
idr0056A	92400	A long noncoding RNA regulates microtubule behaviour during mitosis
idr0056C	13440	A long noncoding RNA regulates microtubule behaviour during mitosis

<https://idr.openmicroscopy.org/about/studies.html>



<https://idr.openmicroscopy.org>

Submitting Data to IDR



To submit data to IDR, email
idr@openmicroscopy.org



Upload original, raw image files.



Fill in metadata templates.



Data published in IDR!

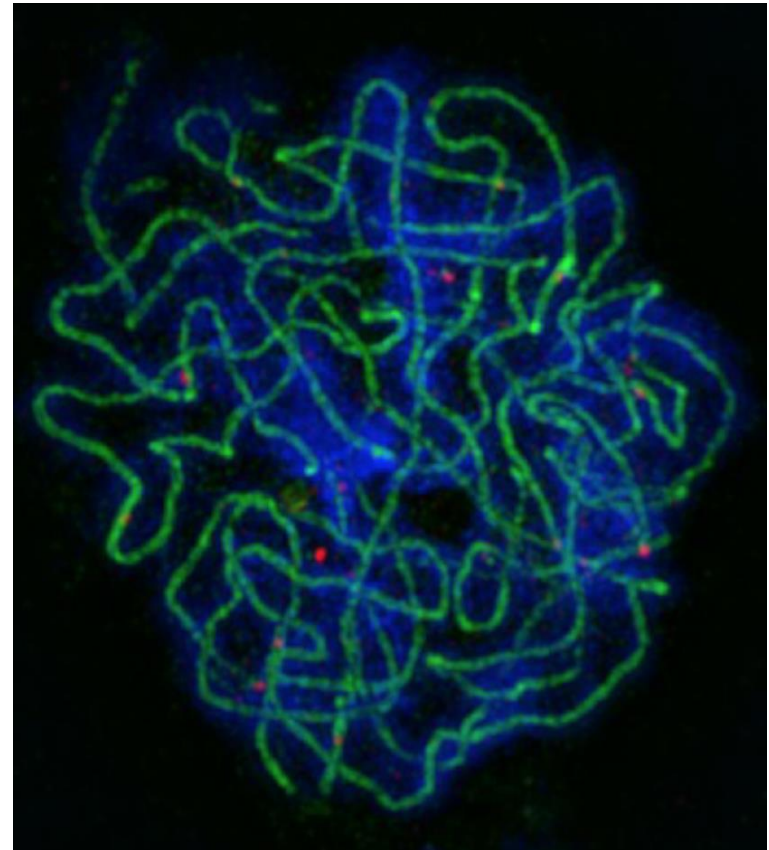


Image from idr0107, Morgan et al.

- ***What data can be submitted to the IDR?***
 - IDR accepts and publishes **cell** and **tissue reference** image datasets.
 - Non-reference cell or tissue image datasets are redirected to BioStudies or the BiImage Archive.
 - If dataset is EM data, it is redirected to [EMPIAR](#).
 - IDR publishes **fully anonymized** data. Full de-identification from Private Health Identifiers (PHI) and Personal Identifying Information (PII) required by submitters before data transfer.

- ***What constitutes an IDR reference dataset?***
 - Datasets **associated** with an existing or upcoming publication.
 - **Complete** datasets – not just images supporting one figure in the publication.
 - Datasets whose metadata can be **integrated** with other datasets via identifiers from well-known biomolecular resources (Ensembl, NCBI Entrez Gene, RefSeq, PubChem, ChEBI etc.).
 - Datasets generated using new imaging **methods** or new analysis methods.
 - Datasets that are likely to be **re-analysed or incorporated** into other studies or integrated with other imaging datasets.

- ***Is there a size limit for data deposition?***
 - There is **no size limit** to the data that can be published in IDR.
 - For very large depositions (>10TB), please contact IDR as early as possible as special arrangements may be needed for data transfer.

- ***How much does it cost to publish data in IDR?***
 - Data deposition and publication is **free-of-charge** to the authors and is supported by grants from the BBSRC, Wellcome Trust and the European Commission.

- ***Which license are datasets published under in IDR?***
 - To to make IDR datasets as widely re-usable as possible, accepted datasets will be published under the Creative Commons Attribution 4.0 International license (**CC BY 4.0**, <https://creativecommons.org/licenses/by/4.0/>) or a more permissive license.



The Image Data Resource (IDR) is a public repository of image datasets published scientific studies, where the community can submit, search and high-quality bio-image data.

[Cell - IDR](#)[Tissue - IDR](#)

Choose search field (optional) ▾

Search for anything...

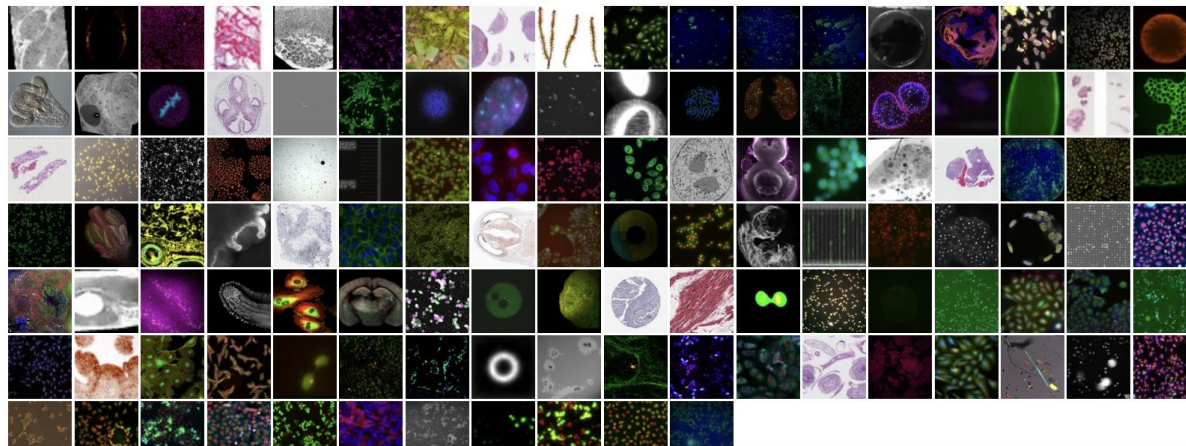
- OVERVIEW
- PUBLISHED STUDIES
- LINKED RESOURCES
- API ACCESS
- DATA DOWNLOAD
- IMAGE TOOLS RESOURCE (ITR)
- ANALYSIS ENVIRONMENTS
- DEPLOYMENT
- FAQ** ←

119 Studies

13,666,786 Images

353 TB

Group Studies by type





IDR Submission Process

Main types of data

- **Images** – upload original raw data, plus analyzed or processed data (if useful).
- **Metadata** – fill in templates to provide information on study and images.

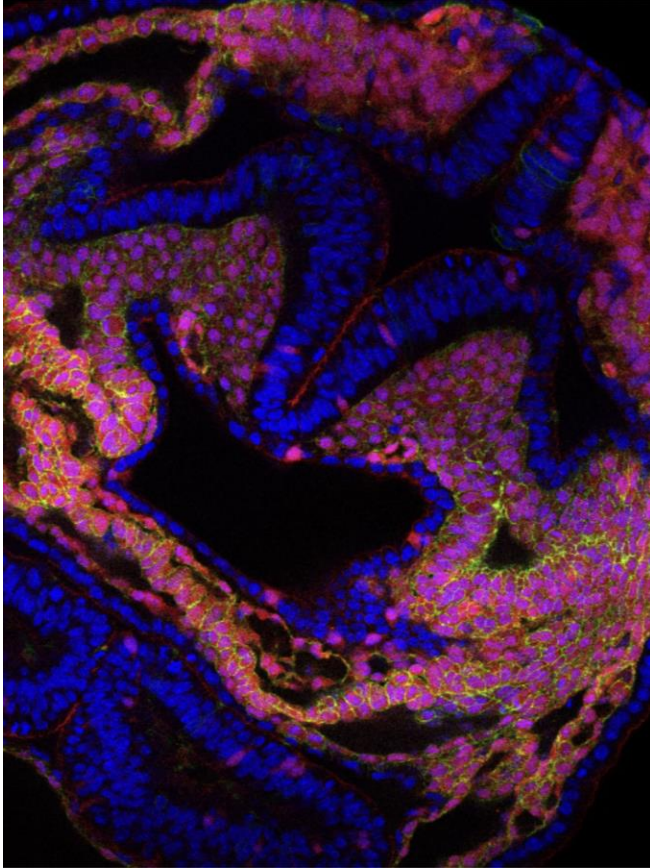


Image from idr0124, Esteban et al.

- Two study types



Screen – high content screens (HCS) or imaging studies performed in a plate format (e.g. 96 well plates).



Experiment – non-screen datasets or studies which group images into a number of distinct experiments.

- Empty templates can be downloaded at <https://github.com/IDR/idr0000-lastname-example/archive/master.zip>
- Examples of completed templates of other studies can be found at <https://github.com/IDR/idr-metadata/>

Metadata Curation in IDR

The screenshot displays the OMERO web interface. The main area shows a grid of microscopy images. The left sidebar contains a file explorer with a list of image files, including 'Exp1_rep1_0min_im1.tif'. The right sidebar shows the metadata for the selected image, 'Exp1_rep1_0min_im1.tif'. The metadata includes acquisition and import dates, dimensions, pixel size, channels, and cell line information.

Experimental metadata

Biomolecular annotations

Analysis results

Category	Value
Image ID	4496763
Owner	Public data
Acquisition Date	2013-11-15 16:02
Import Date	2019-01-11 16:41
Dimensions (XY)	2048 x 2048
Pixel Type	uint16
Pixel Size (XYZ) (µm)	- x - x 0.20
Z-sections/Timepoints	25 x 1
Channels	3-CY5, 5-TMR, 1-DAPI, 7-TRANS
ROI Count	430
Cell Line	BY4741
Gene	YDR536W
Gene Symbol	STL1
Gene	YGR088W
Gene Symbol	CTT1
Dataset Name	Exp1_rep1
Image Name	Exp1_rep1_0min_im1.tif
Term Source 1	4932
REF	
Original Image	
Processed image 1	Exp1_rep1_0min_im1.companion.ome [3D]
Processed image 2	Exp1_rep1_0min_im1.companion.ome [Projection]
Characteristics	BY4741 [Cell Line]

idr0047, Li et al, Scientific Data DOI:10.1038/s41597-019-0106-6



Findable – Original image data cross-linked from publication by DOI



Accessible – Metadata retrievable, open API



Interoperable – Uses common defined vocabularies



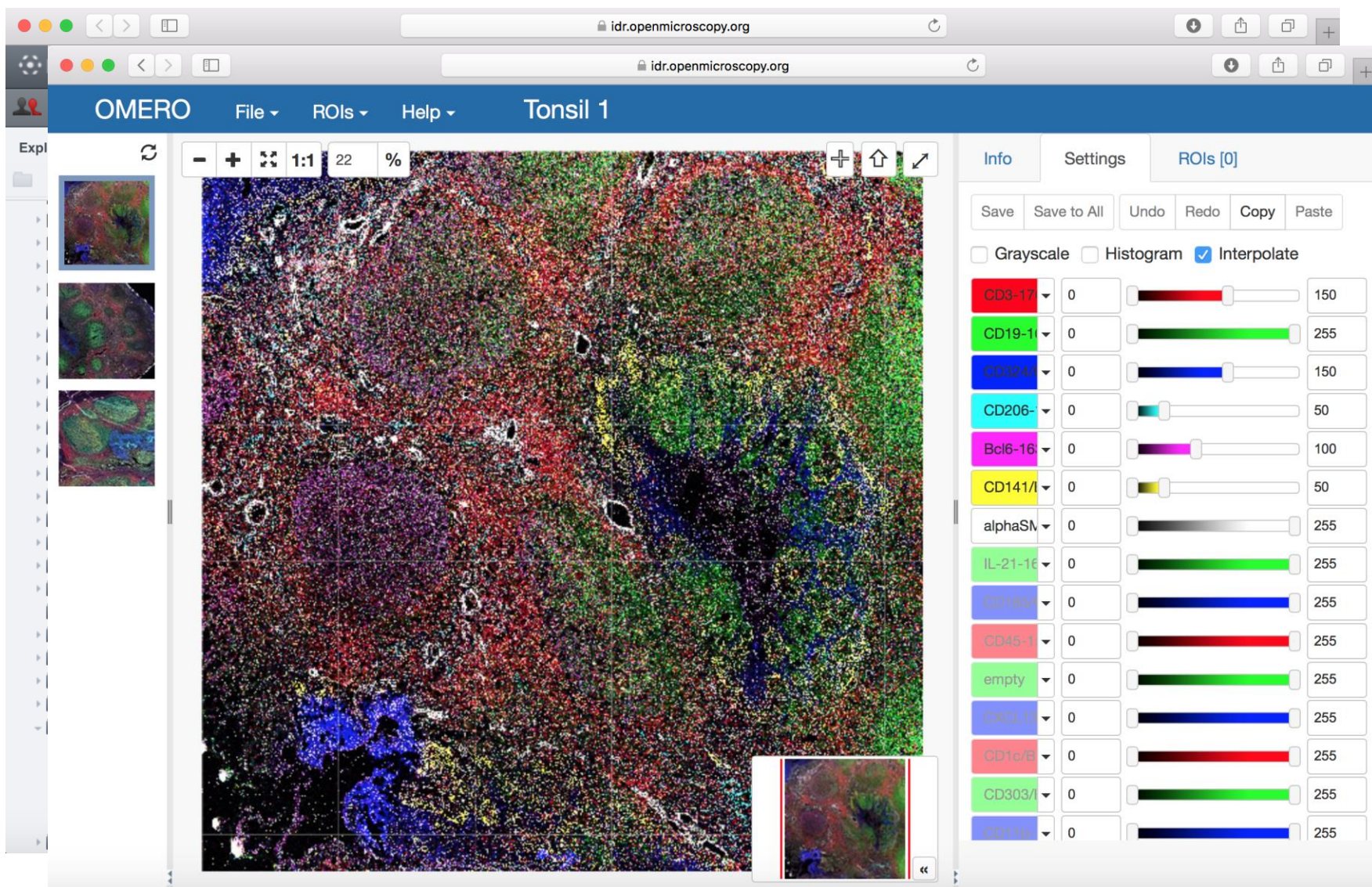
Reusable – Licensed (CC0 or CC BY 4.0), provenance included

Biocuration: to easily find extensive and interlinked information at well-documented, stable resources.

Examples in IDR

- Image Visualisation using IDR Web viewer
- Linked Metadata within IDR
- Link to External Resources
- Querying Metadata within IDR
- Interactive Analysis

Value of Curation – Image Visualisation



idr0054, Durand et al, Scientific Data DOI:10.17867/10000122

Value of Curation – Image Visualisation

The screenshot displays the OMERO web interface for image visualization. The top navigation bar includes links for Studies, Genes, Phenotypes, Cell Lines, siRNAs, Antibodies, Compounds, Organisms, and About. A search bar is located on the right. The main interface is divided into several sections:

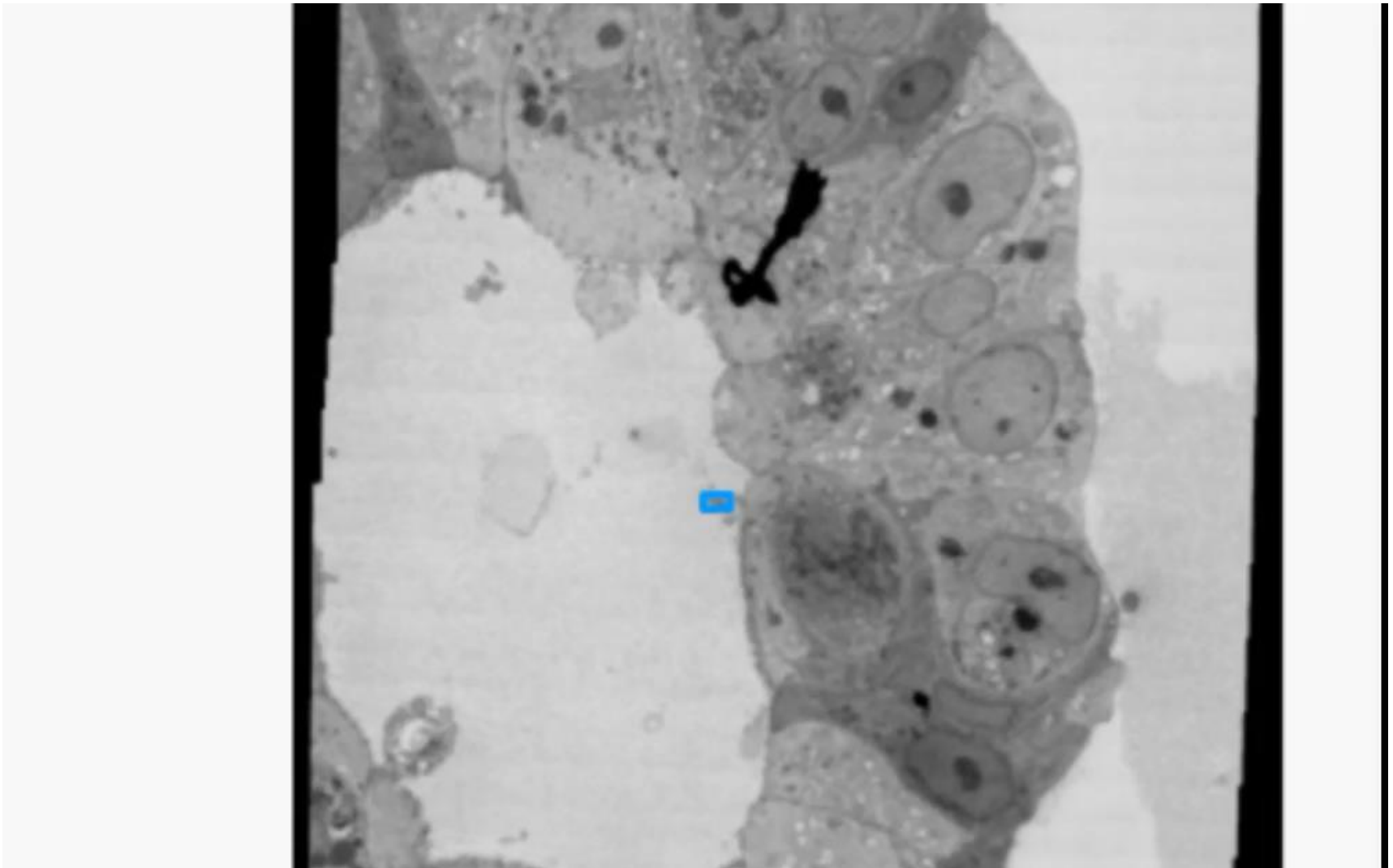
- File Explorer (Left):** Lists various study folders, including 'Intestinal Organoids 2' with sub-items 'hSIOs-1' and 'hSIOs-2'.
- Image Viewer (Center):** Shows a grayscale micrograph of intestinal organoids. A black arrow points to a specific cell. A scale bar at the bottom indicates 50 μm . A small thumbnail of the image is visible in the bottom right corner of the viewer.
- Settings Panel (Right):** Contains tabs for 'Info', 'Settings', and 'ROIs [11]'. It includes controls for 'Save', 'Save to All', 'Undo', 'Redo', 'Copy', and 'Paste'. There are checkboxes for 'Grayscale' (checked), 'Histogram', and 'Interpolate'. A slider for image processing is set between 0 and 10000, with a current value of 3500. Below the slider are buttons for 'Min/Max', 'Full Range', and 'Imported'. A 'User Settings' section shows a thumbnail of the image and the label 'Public data'.

idr0083, Lamers et al *Science* DOI:10.1126/science.abc1669

Value of Curation – Image Visualisation

The screenshot displays the OMERO web interface for image visualization. The main window shows a grayscale electron micrograph of intestinal organoids. The interface includes a navigation menu at the top with options like Studies, Genes, Phenotypes, Cell Lines, siRNAs, Antibodies, Compounds, Organisms, and About. A search bar is located in the top right corner. The left sidebar shows a file browser with a tree view of folders and files, including 'Intestinal Organoids' and 'hSIOs-1'. The central image viewer displays the micrograph with a 500 nm scale bar and a zoom level of 35%. The right sidebar contains a settings panel for the image, including options for Grayscale, Histogram, and Interpolate, and a range slider from 0 to 10000. The settings panel also shows a 'User Settings' section with a thumbnail of the image and a 'Public data' label. The bottom of the interface shows a zoom control and a comments section.

idr0083, Lamers et al *Science* DOI:10.1126/science.abc1669





[EMPIAR home](#) | [Deposition](#) | [REST API](#) | [FAQ](#) | [About EMPIAR](#) | [Policies](#) | [Feedback](#) | [Share](#)

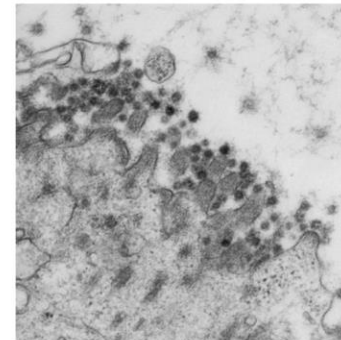
Attributes	1
Publication DOI	
Release Date	
License	
Copyright	
Data Publisher	
Data DOI	
EMPIAR Accession	
Annotation File	

EMPIAR-10404

SARS-CoV-2 productively infects human gut enterocytes

Publication: SARS-CoV-2 productively infects human gut enterocytes
[Lamers MM](#) , [Beumer J](#) , [van der Vaart J](#) , [Knoops K](#) , [Puschhof J](#), [Breugem T](#), [Ravelli RBG](#) , [van Schayck JP](#) , [Mykytyn AZ](#), [Duimel HQ](#), [van Donselaar E](#), [Riesebosch S](#), [Kuijpers HJH](#), [Schipper D](#), [van de Wetering WJ](#), [de Graaf M](#), [Koopmans M](#) , [Cuppen E](#) , [Peters PJ](#), [Haagmans B](#) , [Clevers H](#)
Science
[DOI: 10.1126/science.abc1669](https://doi.org/10.1126/science.abc1669)

Contains:
stitched maps



Related IDR entry: [idr0083](#)
Deposited: 2020-04-30
Released: 2020-05-01
Last modified: 2020-05-01
Dataset size: 66.5 GB
Dataset DOI: [10.6019/EMPIAR-10404](https://doi.org/10.6019/EMPIAR-10404)
Experimental metadata: [Download xml](#)

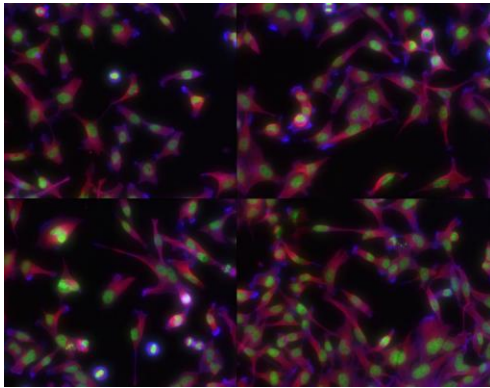
Quick links

- [EMDB](#)
- [PDBe](#)
- [BioImage Archive](#)
- [EMPIAR Quick tour](#)
- [Statistics](#)
- [Re-use case study](#)
- [EMPIAR@PDBj](#)

EMPIAR citations

[Cryo-EM structure of the potassium-chloride cotransporter KCC4 in lipid nanodiscs.](#)
Reid MS, Kern DM, Brohawn SG. (2020)
[Development of basic building blocks](#)

Value of Curation - Linked Metadata



idr0012

OMERO Studies Genes Phenotypes Cell Lines siRNAs Antibodies

Public

Type Gene Symbol... Match case Add filter

- Gene 1
 - ASH2L (326) 6
 - idr0022-koedoot-cellmigration/screenA (288) 8
 - idr0006-fong-nuclearbodies/screenA (16) 1
 - idr0009-simpson-secretion/screenA (12) 6
 - idr0013-neumann-mitochek/screenA (6) 3
 - idr0012-fuchs-cellmorph/screenA (2) 1
 - HT28 2
 - HT28 [Well G13, Field 1]
 - HT28 [Well G13, Field 2]
 - idr0010-doil-dnamage/screenA (2) 1

Attributes 8

Cell Lines

Added by: Public data

Cell Line	HeLa
-----------	------

Gene

Added by: Public data

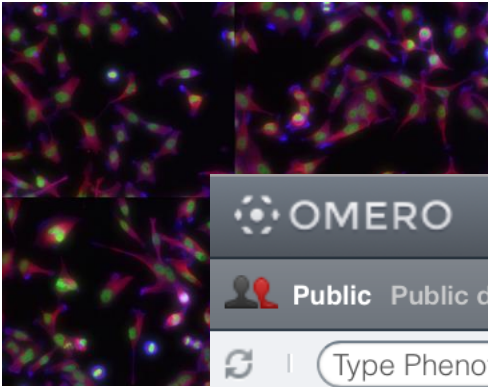
Gene Identifier	9070
Gene Symbol	ASH2L

Phenotype

Added by: Public data

Phenotype	elongated cells
Phenotype Term Name	elongated cell phenotype
Phenotype Term Accession	CMPO_0000077

Value of Curation - Linked Metadata



Attributes 8

Cell Lines

Added by: Public data

Cell Line

HeLa

OMERO

Studies

Genes

Phenotypes

Cell Lines



Public

Public data



Type Phenotype...



Match case



Phenotype 1



CMPO_000077 (20872) 8



idr0028-pascualvargas-rhogtpases/screenD (3920) 4



idr0028-pascualvargas-rhogtpases/screenA (3808) 4



idr0028-pascualvargas-rhogtpases/screenC (3220) 4



idr0033-rohban-pathways/screenA (2916) 12



idr0028-pascualvargas-rhogtpases/screenB (2856) 4



idr0001-graml-sysgro/screenA (2788) 122



idr0008-rohn-actinome/screenB (1098) 10



idr0012-fuchs-cellmorph/screenA (266) 55

9070



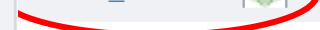
ASH2L



elongated cells

elongated cell phenotype

CMPO_000077



OMERO

Studies

Genes

Phenotypes



Public



Type Gene Symbol...



Gene 1



ASH2L (326) 6



idr0022-koedoot-cellmigration/screenA



idr0006-fong-nuclearbodies/screenA (1)



idr0009-simpson-secretion/screenA (12)



idr0013-neumann-mitochek/screenA (6) 3



idr0012-fuchs-cellmorph/screenA (2) 1



HT28 2



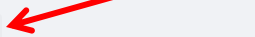
HT28 [Well G13, Field 1]



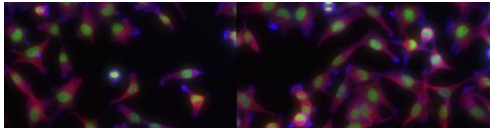
HT28 [Well G13, Field 2]



idr0010-doil-dnadamage/screenA (2) 1



Value of Curation - Link to External Resources



Log in

Gene Advanced

Full Report ▾

Send to: ▾

Hide sidebar >>

Table of contents

- Summary
- Genomic context
- Genomic regions, transcripts, and products
- Expression
- Bibliography
- Phenotypes
- Variation
- Pathways from PubChem
- Interactions
- General gene information
 - Markers, Related pseudogene(s), Homology, Gene Ontology
- General protein information
- NCBI Reference Sequences (RefSeq)
- Related sequences
- Additional links

Download Datasets

Summary

Official Symbol ASH2L provided by HGNC
Official Full Name ASH2 like, histone lysine methyltransferase complex subunit provided by HGNC
Primary source [HGNC:HGNC:744](#)
See related [Ensembl:ENSG00000129691](#) [MIM:604782](#); [AllianceGenome:HGNC:744](#)
Gene type protein coding
RefSeq status VALIDATED
Organism [Homo sapiens](#)
Lineage Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo
Also known as ASH2; Bre2; ASH2L1; ASH2L2
Summary Enables beta-catenin binding activity and transcription cis-regulatory region binding activity. Contributes to histone methyltransferase activity (H3-K4 specific). Involved in histone H3-K4 methylation; positive regulation of cell population proliferation; and response to estrogen. Acts upstream of or within cellular response to DNA damage stimulus. Located in nucleus. Part of MLL3/4 complex and Set1C/COMPASS complex. [provided by Alliance of Genome Resources, Apr 2022]
Expression Ubiquitous expression in testis (RPKM 20.9), fat (RPKM 13.3) and 25 other tissues [See more](#)
Orthologs [mouse](#) [all](#)

Attributes 8

Cell Lines

Added by: Public data

HeLa

Public data

9070 

ASH2L

Public data

elongated cells

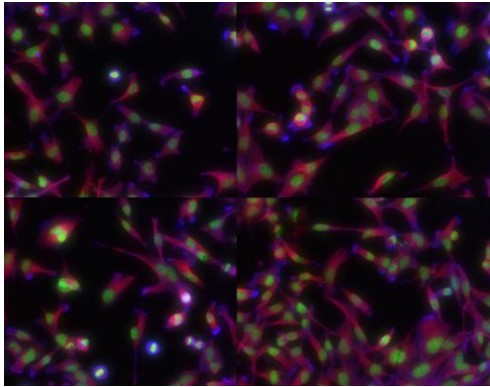
Name

elongated cell phenotype

Accession

CMPO_0000077 

Value of Curation - Link to External Resources



idr0012

Attributes 8

Cell Lines

Added by: Public data

Cell Line	HeLa
-----------	------

Gene

Added by: Public data

Gene Identifier	9070
Gene Symbol	ASH2L

Phenotype

Added by: Public data

Search CMPO Search

phenotype
7 ←

elongated cell phenotype

http://www.ebi.ac.uk/cmipo/CMPO_0000077 Copy

A phenotype observation at the level of the cell shape where the cell is elongated, with a length notably greater than its width

Tree view

Graph view
Reset tree
Show all siblings

- cellular phenotype
 - cellular component phenotype
 - cell component morphology phenotype
 - cell component shape phenotype
 - elongated cell phenotype**
 - cell morphology phenotype
 - elongated cell phenotype**
 - single cell phenotype
 - cell morphology phenotype
 - elongated cell phenotype**

Term information

Subsets

- cmipo

has broad synonym

- bipolar cell

id

- CMPO:0000077

prefLabel

- elongated cell

Organism	NCBI Taxonomy https://www.ncbi.nlm.nih.gov/taxonomy
Study Type Screen Type (HCS) Screen Technology Type (HCS) Library Type (HCS) Protocol	Experimental Factor Ontology (EFO) https://www.ebi.ac.uk/efo/
Imaging Method	Biological Imaging Methods Ontology (FBbi) https://www.ebi.ac.uk/ols/ontologies/fbbi
Phenotype	Cellular Microscopy Phenotype Ontology (CMPO) https://www.ebi.ac.uk/cmppo/
Gene	Ensembl https://www.ensembl.org/ NCBI Gene https://www.ncbi.nlm.nih.gov/gene/
Protein	UniProt https://www.uniprot.org/
Clinical/Pathology	SNOMED CT https://bioportal.bioontology.org/ontologies/SNOMEDCT
Compound	PubChem https://pubchem.ncbi.nlm.nih.gov/

Value of Curation – Querying Metadata

◀ Home

Attribute	Operator	Value
Compound Name ▾	contains ▾	Remdesivir

Search **Compound Name** contains **Remdesivir** found **216** images in **1** screen

ID	Images	Publication Title
▼ idr0094B	216	A SARS-CoV-2 cytopathicity dataset generated by high-content screening of a larg...

SARS-CoV-2 and Remdesivir

The screenshot displays the IDR web application interface. At the top, there is a navigation bar with tabs for Studies, Genes, Phenotypes, Cell Lines, siRNAs, Antibodies, Compounds, Organisms, and About. A search bar is located on the right side of the top bar. Below the navigation bar, the main content area is divided into a left sidebar, a central grid, and a right-hand panel.

Left Sidebar: Shows a tree view of the data structure. The selected path is: Compound 1 > Remdesivir (216) 1 > idr0094-ellinger-sarscov2/screenB (216) 3 > ESP0025721 72. A red arrow points from the text "20µM" to the top-left image in the grid.

Central Grid: A 10x10 grid of microscopy images showing the effect of Remdesivir on SARS-CoV-2. The images show a concentration gradient from top-left to bottom-right. A red arrow points from the text "0.0064µM" to the bottom-right image in the grid.

Right-hand Panel: Contains metadata for the selected compound, ESP0025721. The metadata includes: Plate ID: 7824, Owner: Public data, Creation Date: 2020-09-30 18:38:55, and various counts for Attributes, Attachments, Comments, Tags, Ratings, and Others.

Value of Curation – Interactive Analysis

The screenshot displays the IDR interface. On the left, a tree view shows the dataset structure: **idr0094-ellinger-sarscov2** (Public data) containing **screenA 64** and **screenB 102**. Under **screenB**, a list of 30 ESP IDs is shown, ranging from ESP0025712 to ESP0025739. The right panel provides details for the selected screen:

- General** tab selected.
- Owner:** Public data (Show all)
- Screen Details:** Publication Title: A SARS-CoV-2 cytopathicity dataset generated by high-content screening of a large drug repurposing collection. Screen Description: Dose response screen.
- Creation Date:** 2020-09-30 15:09:06
- Plate Count:** 102 plates
- Attributes:** 3
 - Data Publisher:** Universil
 - Data DOI:** <https://doi.org/10.17867/10001700148b>
 - BioStudies Accession:** S-BIAD29 <https://www.ebi.ac.uk/biostudies/studies/S-BIAD29>
 - Annotation File:** [idr0094-screenB-annotation.csv](https://github.com/IDR/idr0094-ellinger-sarscov2/blob/HEAD/screenB/idr0094-screenB-annotation.csv) <https://github.com/IDR/idr0094-ellinger-sarscov2/blob/HEAD/screenB/idr0094-screenB-annotation.csv>
- Analysis Notebook:** Added by: Public data. Study Notebook: [idr0094-ic50.ipynb](#) (indicated by a red arrow).
- Attachments:** 1
- Comments:** 0
- Tags:** 0
- Ratings:** 0

idr0094, Ellinger et al *Sci Data* DOI:10.1038/s41597-021-00848-4

IC50 exploration

This notebook demonstrates how to process plates associated to the paper [A SARS-CoV-2 cytopathicity dataset generated by high-content screening of a large drug repurposing collection](#). We explore the metadata linked to [idr0094-ellinger-sarscov2/screenB](#).

Calculate the half maximal inhibitory concentration IC50

The half maximal inhibitory concentration (IC50) is a measure of the potency of a substance in inhibiting a specific biological or biochemical function. IC50 is a quantitative measure that indicates how much of a particular inhibitory substance (e.g. drug) is needed to inhibit, in vitro, a given biological process or biological component by 50%.

```
In [11]: calculate_IC50 <- function(data){
  IC50 <- NA
  data.LL.4 <- NULL
  ctest <- filter(data, Inhibition != "NaN")
  data.LL.4 <- drm(Inhibition ~ Concentration, data = ctest, fct = LL.4(), control = drmc(errorrm=FALSE))
  if (is.null(data.LL.4$convergence)) {
    IC50 <- ED(data.LL.4, 50, interval = "delta")[1]
  }
  returned_values <- list("ic50" = IC50, "data" = data.LL.4)
}
```

```
In [12]: values <- calculate_IC50(data)
IC50 <- values$ic50
IC50
```

Estimated effective doses

	Estimate	Std. Error	Lower	Upper
e:1:50	0.80772	0.12019	0.55701	1.05842
	0.807715887041243			

Plot the Dose Response curve

```
In [13]: options(repr.plot.width=6, repr.plot.height=5.8)
plot(values$data, broken = TRUE, type = "all",
      main = "Dose Response Curve (DRC)", xlim = c(0, 100),
      xlab = "Log Concentration (uM)",
      ylab = "Percent Inhibition (percent)")
abline(h = 50, col = "cyan")
abline(v = IC50, col = "green")
```

Discussion

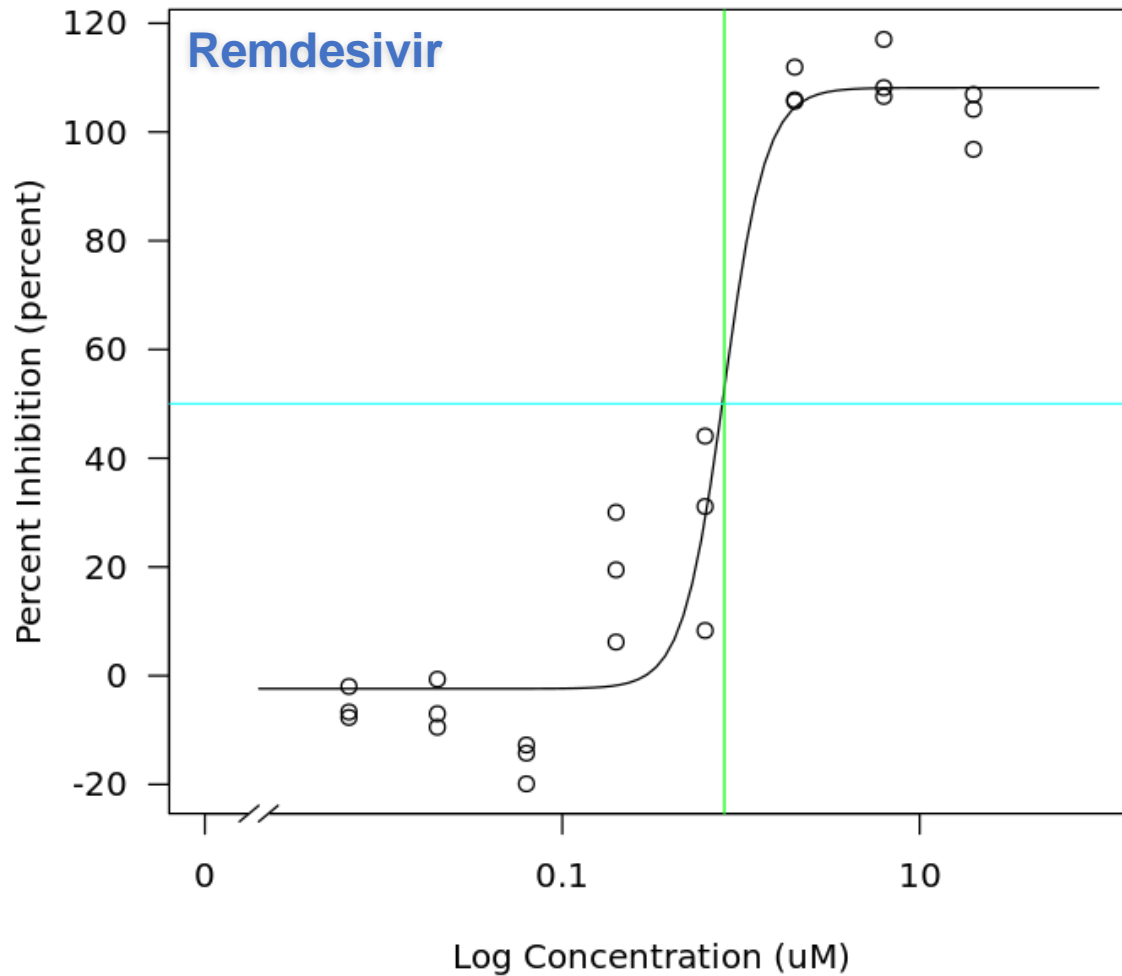
The activity of the reference compound, remdesivir (IC50 = 0.76 μ M) was confirmed in this study.





Remdesivir targets the viral nsp12 RNA-dependent RNA polymerase⁽⁸⁾ and is currently under evaluation in an adaptive, randomized, double-blind, placebo-controlled phase III clinical trial⁽⁹⁾.

Dose Response Curve of Remdesivir



Dose Response Curve (DRC)



 WorkflowHub  Browse ▾  Create ▾  Help ▾

 **Calculate the half maximal inhibitory concentration for each compound used in SARS-CoV-2 investigation** Version 2 (latest) ▾

Workflow Type: Jupyter

Stable



<https://workflowhub.eu/workflows/238>

The Image Data Resource (IDR) is a public repository of image datasets from published scientific studies, where the community can submit, search and access high-quality bio-image data.

Cell - IDR Tissue - IDR

Choose search field (optional) Search for anything...

119 Studies 13,666,786 Images 353 TB

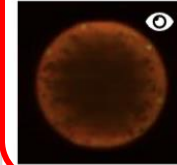
Group Studies by type

119 Studies

13,666,786 Images

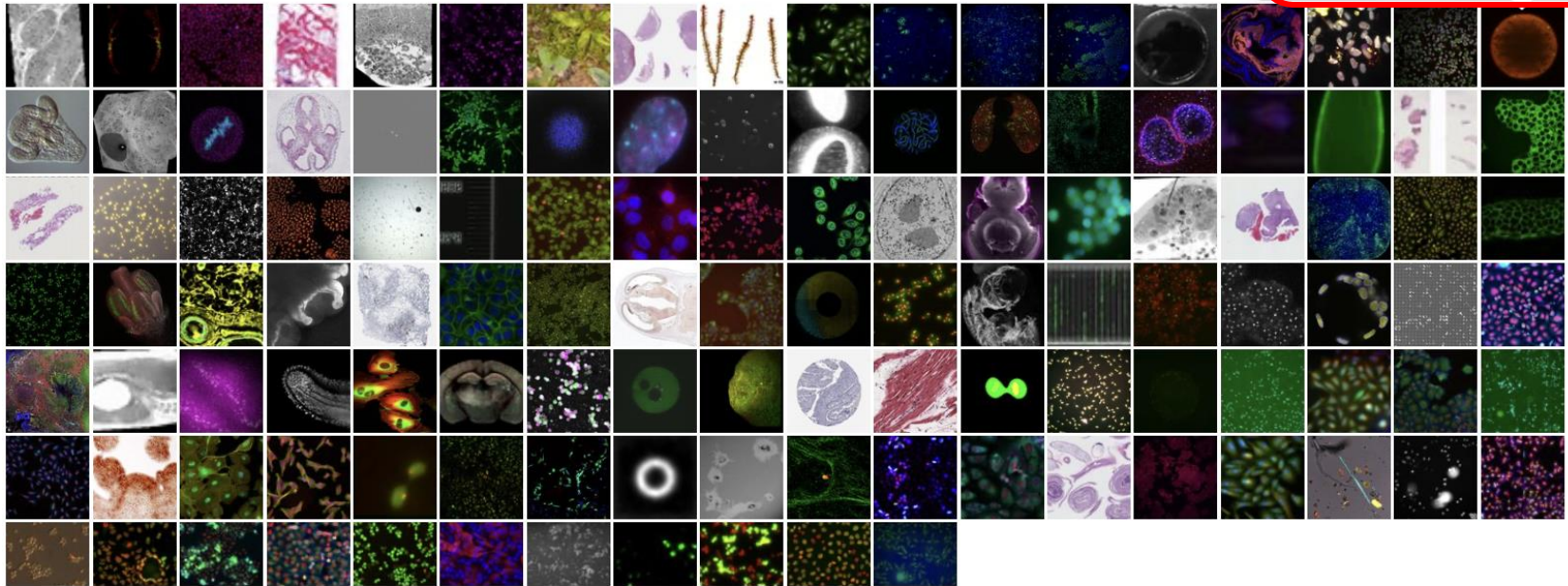
idr0118

Keenan SE et. al

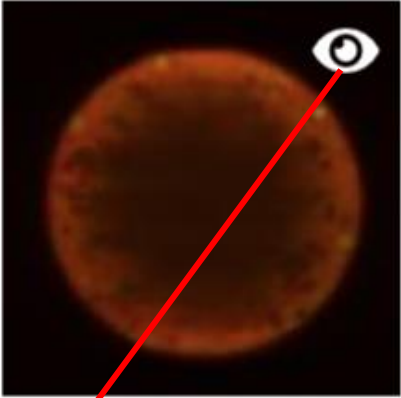


1 Experiment **20 Images**
Dynamics of Drosophila
endoderm specification

Group Studies by type



idr0118 **Keenan SE et. al**



1 Experiment
Dynamics of Drosophila
endoderm specification

Image viewer

Study data

Publication (PubMed)

The Image Data Resource (IDR) is a public repository of image datasets from published scientific studies, where the community can submit, search and access high-quality bio-image data.

Cell - IDR Tissue - IDR

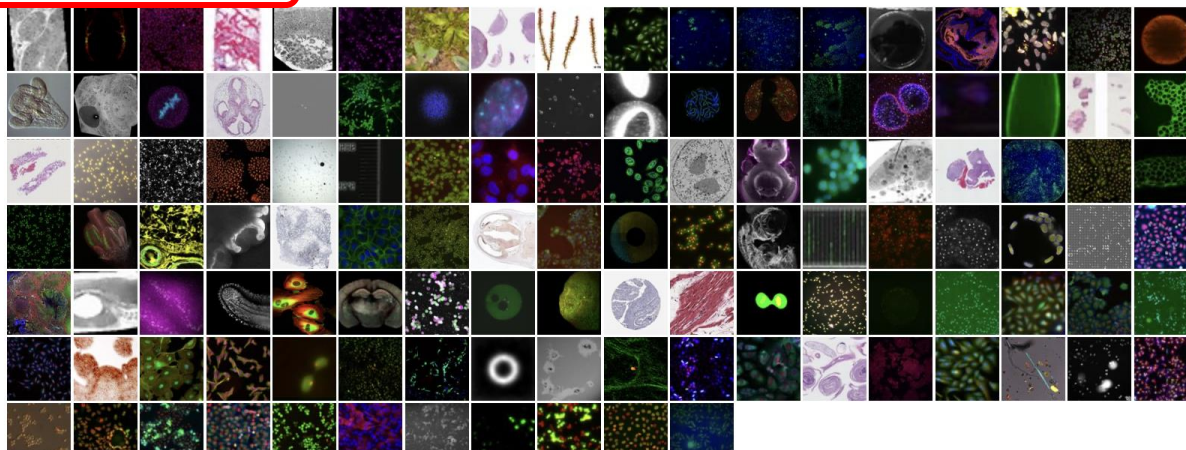
Choose search field (optional) Search for anything...

119 Studies

13,666,786 Images

353 TB

Group Studies by type



The Image Data Resource (IDR) is a public repository of image datasets from published scientific studies, where the community can submit, search and access high-quality bio-image data.

Cell - IDR Tissue - IDR

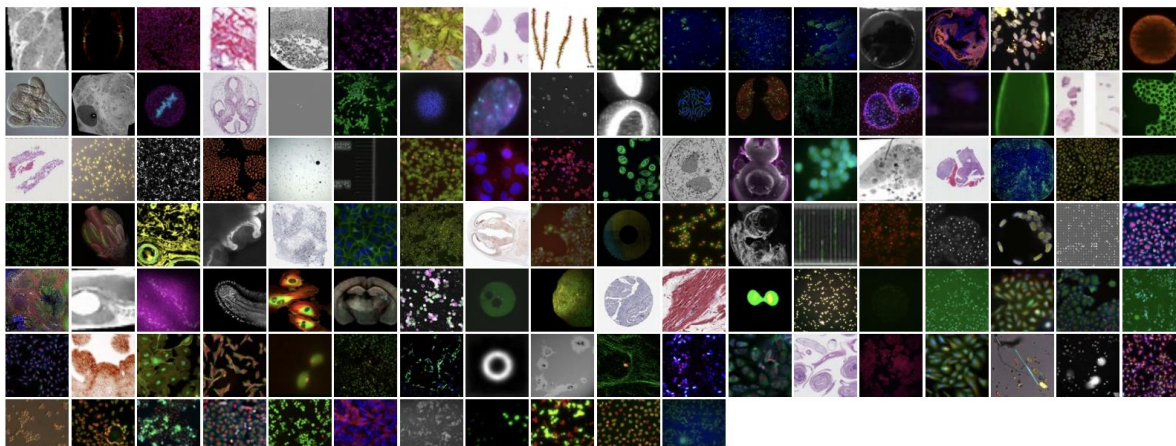
Choose search field (optional) Search for anything...

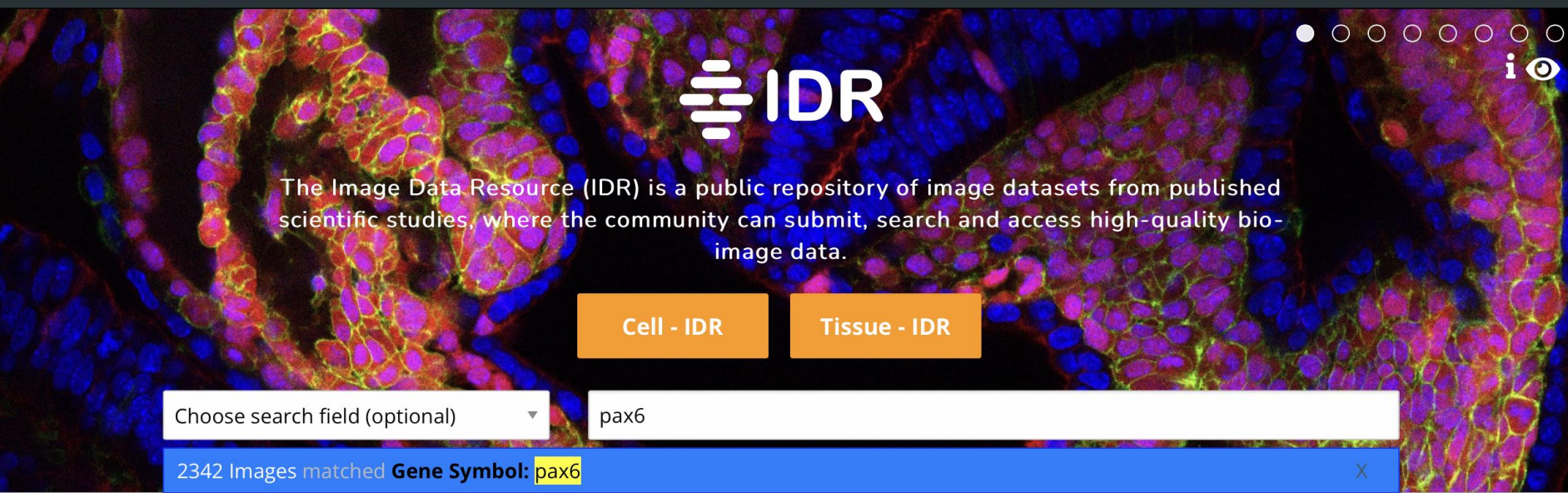
119 Studies

13,666,786 Images

353 TB

Group Studies by type





The Image Data Resource (IDR) is a public repository of image datasets from published scientific studies, where the community can submit, search and access high-quality bio-image data.

Cell - IDR

Tissue - IDR

Choose search field (optional) ▾ pax6

- 2342 Images matched **Gene Symbol:** pax6
- 991 Images matched **Antibody:** pax6
- 9 Images matched **HGNC Gene Symbol:** pax6
- 6 Images matched **Nearest Human Homologue:** pax6
- 4 Images matched **Original Gene Target:** pax6
- 3 Images matched **Gene Symbol Synonyms:** eyeless tumor-head-63 l(4)102cdh ey eyel pax-6 ok107 ey/pax6 eye l(4)33 pax6 ey dpax-6 l(4)102cdr pax6 cg1464

◀ Home

Attribute: Gene Symbol Operator: equals Value: pax6

Search **Gene Symbol** equals **pax6** found **2342** images in **9** experiments/screens

ID	Images	Publication Title
▶ idr0043A	1027	Proteomics. Tissue-based map of the human proteome.
▶ idr0070A	855	The HUDSEN Atlas: a three-dimensional (3D) spatial framework for studying gene ...
▶ idr0022A	288	Uncovering the signaling landscape controlling breast cancer cell migration identi...
▶ idr0114A	136	Enabling research with human embryonic and fetal tissue resources
▶ idr0009A	14	Genome-wide RNAi screening identifies human proteins with a regulatory functio...
▶ idr0013A	9	Phenotypic profiling of the human genome by time-lapse microscopy reveals cell ...
▶ idr0093A	9	High content genome-wide siRNA screen to investigate the coordination of cell siz...
▶ idr0010A	2	RNF168 binds and amplifies ubiquitin conjugates on damaged chromosomes to a...
▶ idr0012A	2	Clustering phenotype populations by genome-wide RNAi and multiparametric im...

Browser address bar: idr.openmicroscopy.org

Navigation: [Home](#) | [CELL - IDR](#) | [TISSUE - IDR](#) | [ABOUT](#) | [SUBMISSIONS](#)

Search criteria: Attribute: Gene Symbol, Operator: equals, Value: pax6

Search Gene Symbol equals pax6 found 2342 images in 9 experiments/screens

ID	Images	Publication Title
idr0043A	1027	Proteomics. Tissue-based map of the human proteome.
idr0070A	855	The HUDSEN Atlas: a three-dimensional (3D) spatial framework for studying gene ...
idr0022A	288	Uncovering the signaling landscape controlling breast cancer cell migration identi...
idr0114A	136	Enabling research with human embryonic and fetal tissue resources

Browser address bar: idr.openmicroscopy.org

Navigation: [Home](#) | [CELL - IDR](#) | [TISSUE - IDR](#) | [ABOUT](#) | [SUBMISSIONS](#)

Search criteria: Attribute: **Gene Symbol** | Operator: **equals** | Value: **pax6**

Search **Gene Symbol** equals **pax6** found **2342** images in **9** experiments/screens

ID	Images	Publication Title
idr0043A	1027	Proteomics. Tissue-based map of the human proteome.
idr0070A	855	The HUDSEN Atlas: a three-dimensional (3D) spatial framework for studying gene ...
idr0022A	288	Uncovering the signaling landscape controlling breast cancer cell migration identi...
idr0114A	136	Enabling research with human embryonic and fetal tissue resources

IDR Search

Attribute	Operator	Value
Gene Symbol	equals	pax6

- Any
- Study
 - Imaging Method
 - License
 - Name (IDR number)
 - Organism
 - Publication Authors
 - Publication Title
 - Screen Technology Type
 - Screen Type
 - Study Type
- Image
 - Antibody
 - Antibody Identifier
 - Cell Line
 - Compound Name
 - Gene Identifier
 - ✓ Gene Symbol
 - InChIKey
 - Pathology
 - Pathology Identifier
 - Phenotype
 - Phenotype Term Accession
 - Protein
 - PubChem CID
 - siRNA Identifier
 - siRNA Pool Identifier

- contains
- ✓ equals

Browser address bar: idr.openmicroscopy.org

Navigation: **IDR** CELL - IDR TISSUE - IDR ABOUT SUBMISSIONS

Home button (indicated by a red arrow)

Search criteria: Attribute: Gene Symbol, Operator: equals, Value: pax6

Search Gene Symbol equals pax6 found 2342 images in 9 experiments/screens

ID	Images	Publication Title
▶ idr0043A	1027	Proteomics. Tissue-based map of the human proteome.
▼ idr0070A	855	The HUDSEN Atlas: a three-dimensional (3D) spatial framework for studying gene ...
▶ idr0022A	288	Uncovering the signaling landscape controlling breast cancer cell migration identi...
▶ idr0114A	136	Enabling research with human embryonic and fetal tissue resources

Exploring IDR

Analyze Data



From publication to IDR

Find a specific IDR study from a publication. View images and Regions of Interest (ROIs) in IDR. Explore study metadata.



From gene to phenotypes

Query images by gene. Explore retrieved images and metadata, including associated phenotypes. View images in context of a plate.



From compound to analytics

Query images by compound. Find SARS-CoV-2 dataset. Explore images in relation to analytical data submitted by authors.

Exploring IDR

Analyze Data

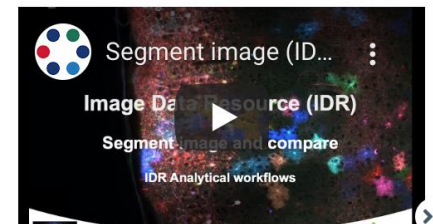
The IDR server is built with **OMERO**, allowing access to all image data and metadata via an open API in Python, R, Java, MATLAB and REST/JSON. See the **OMERO API guide** for more information.

For examples of analysis tools working with OMERO to access and analyze data, see the **analysis tools guide**.



Analysis environment setup

Set-up local or on-the-cloud analysis environment. Access images and metadata using API. Run analysis examples. Add your own code.



Segment image and compare

Fetch image and labels from IDR. Segment nuclei using **StarDist**. Compare new labels with original. Validate software package. Save locally segmentation polygons as geojson.

Which diabetes related genes are expressed in the pancreas?

HumanMine
v12 February 2022
An integrated database of *Homo sapiens* genomic data

<https://www.humanmine.org/humanmine>

```
TISSUE = "Pancreas"  
DISEASE = "diabetes"
```

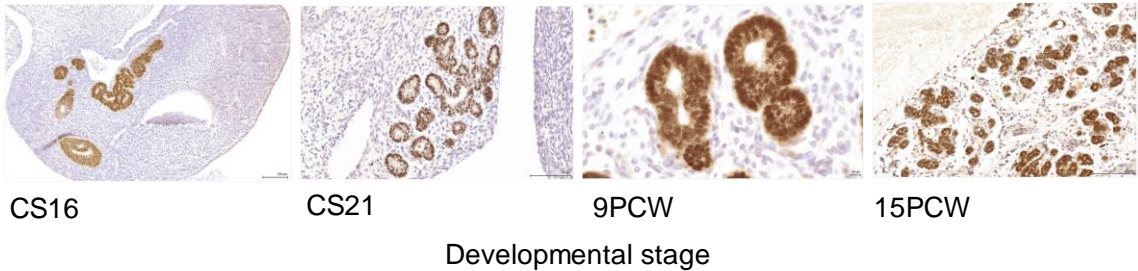
```
query.add_constraint("proteinAtlasExpression.tissue.name", "=", TISSUE)  
query.add_constraint("proteinAtlasExpression.level", "GNE OF", ("Medium", "High"))  
query.add_constraint("organism.name", "=", "Homo sapiens")  
query.add_constraint("diseases.name", "CONTAINS", DISEASE)  
  
<BinaryConstraint: Gene.diseases.name CONTAINS diabetes>  
  
Collect the genes  
  
upin_tissue = list()  
for row in query.rows():  
    upin_tissue.append(row["symbol"])  
unique = set(upin_tissue)  
genes = sorted(genes, reverse=True)
```

Genes found

WFS1 VEGFA TCF7L2 TBC1D4 SOD2 SLC30A8 PTPN22 **PDX1**
MIA3 KCNJ11 IRS2 IRS1 INSR INS IGF2BP2 IER3IP1
HNF4A HNF1B HMG1A HPE GPD2 GCK ENPP1 EIF2AK3
DNAJC3 CEL CAPN10 APPL1 AKT2 ABCC8

IDR MULTIOMICS API

Images linked to gene PDX1



Search for images in IDR associated to the genes found in Humanmine

From the list of genes found using the intermine API, we are now looking in [Image Data Resource](#) for studies linked to those genes and with tissue as a Sample Type.

```
TYPE = "gene"  
SAMPLE_TYPE = "tissue"  
EXPRESSION_KEY = "Expression Pattern Description"  
EXPRESSION = "Islets" # "Brain"  
KEYS = ("phenotype":  
    {"phenotype":  
        "Phenotype Term Name",  
        "Phenotype Term Accession",  
        "Phenotype Term Accession URL",  
    })  
  
projects = list()  
for gene in genes:  
    qs1 = {'key': TYPE, 'value': gene}  
    url1 = URL.format(**qs1)  
    json = session.get(url1).json()  
    for m in json['maps']:  
        qs2 = {'key': TYPE, 'value': gene}  
        url2 = SCREENS_PROJECTS_URL.format(**qs2)  
        json = session.get(url2).json()  
        for p in json['projects']:  
            value = find_type("project", p['id'])  
            if value > -1:  
                projects.append(value)
```


Diabetes related genes expressed in pancreas Version 1

Overview Files Related items

Workflow Type: Jupyter

Stable

Summary

This notebook shows how to integrate genomic and image data resources. This notebook looks at the question **Which diabetes related genes are expressed in the pancreas?**

Steps:

- Query humanmine.org, an integrated database of Homo sapiens genomic data using the intermine API to find the genes.
- Using the list of found genes, search in the Image Data Resource (IDR) for images linked to the genes, tissue and disease.
-

We use the intermine API and the IDR API

The notebook can be launched in [My Binder](#)

Inputs

Parameters needed to configure the workflow:

Screenshot

Creators and Submitter

Creator

Jean-Marie Burel

Submitter

Jean-Marie Burel

Discussion Channel

[Launch on MyBinder](#)

Citation

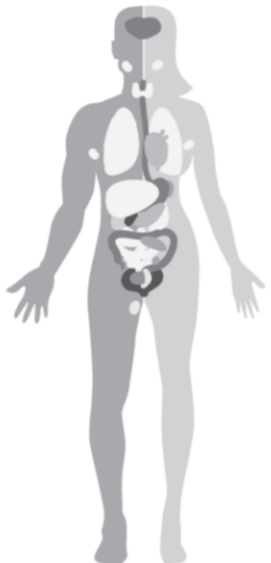
Copy

Burel, J.-M. (2021). *Diabetes related genes expressed in pancreas*. WorkflowHub. <https://doi.org/10.48546/WORKFLO>

IDR and Human Protein Atlas

THE HUMAN PROTEIN ATLAS

MENU HELP NEWS



RNA expression (TPM)¹ Protein expression (score)²

Brain

Endocrine tissues

Bone marrow & immune system

Muscle tissues

Public

Explore Tags Shares Add filter

- 125920_A_7_1.tif
- 125920_A_7_2.tif
- 125920_A_7_4.tif
- 125920_A_7_5.tif
- 125920_A_7_6.tif
- 125920_A_7_8.tif
- 125920_A_8_1.tif
- 125920_A_8_2.tif
- 125920_A_8_3.tif
- 125920_A_8_4.tif
- 125920_A_8_5.tif
- 125920_A_8_6.tif
- 125920_A_8_7.tif
- 125920_A_9_1.tif
- 125920_A_9_2.tif
- 125920_A_9_3.tif
- 125920_A_9_4.tif
- 125920_A_9_5.tif
- 125920_A_9_6.tif
- 125920_A_9_7.tif
- 125920_A_9_8.tif
- 125920_B_1_1.tif
- 125920_B_1_3.tif
- 125920_B_1_5.tif
- 125920_B_1_6.tif
- 125920_B_1_7.tif
- 125920_B_2_1.tif

OMERO File - ROIs - Help - 125920_A_7_8.tif

Dimensions (XY): 3000 x 3000
Pixels Type: uint8
Pixels Size (XYZ) (µm): -
Z-sections/Timepoints: 1 x 1
Channels: 0, 1, 2
ROI Count: 0

Attributes 6

Antibody

Added by: Public data

Antibody Identifier: CAB058686

Antibody supplementary

Added by: Public data

Dilution Factor: 1:1000
Retrieval Method: HIER pH6
Human Protein Atlas version: v18

Gene

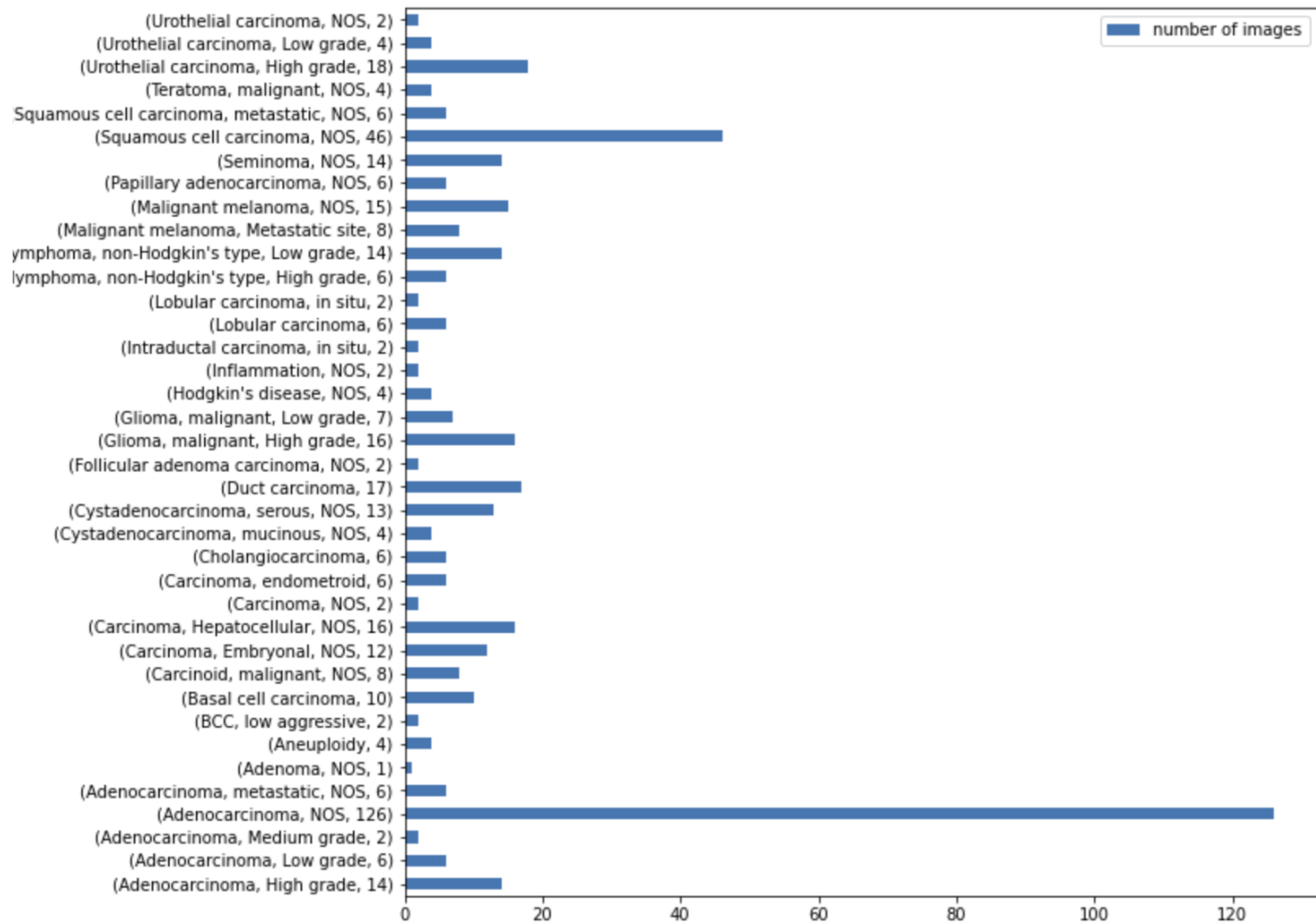
Added by: Public data

Gene Identifier: ENSG00000148516

Tables

Human Protein Atlas:

Number of images associated to the gene PDX1 with abnormal pathology status



In Situ Sequencing - Linking Multimodal Metadata

The screenshot displays the OMERO web interface for a dataset named 'cell002_processed'. The main view shows a microscopy image of a cell with numerous small, colored spots (red, green, yellow, blue, purple) and larger, overlapping circles of various colors. A scale bar at the bottom left indicates 5 μm. The interface includes a top navigation bar with links for Studies, Genes, Phenotypes, Cell Lines, siRNAs, Antibodies, Compounds, Organisms, and About. A search bar is located in the top right. The left sidebar shows a file explorer with folders for 'idr0101' and 'Fibroblasts'. The right sidebar contains settings for ROIs, including options for Grayscale, Histogram, and Interpolate, and a list of channels with their respective values and ranges. The bottom of the interface shows a zoom slider and a comments section.

idr0101, Payne et al *Science* DOI:10.17867/10000169

In Situ Sequencing - Linking Multimodal Metadata

IDR Studies Genes Phenotypes Cell Lines siRNAs Antibodies Compounds Organisms About

Public Public data

Explore Tags Shares

idr0101-payne-insitugenomeseq

Add filter

General Acquisition Preview

Term Source 1 NCBITaxon
REF:
Term Source 1 NCBITaxon_9606
Accession:
Characteristics GM23248
[Cell Line]:

bulk_annotations

Download as CSV: [Whole Table](#)

Show current page as: [CSV](#) | [JSON](#)

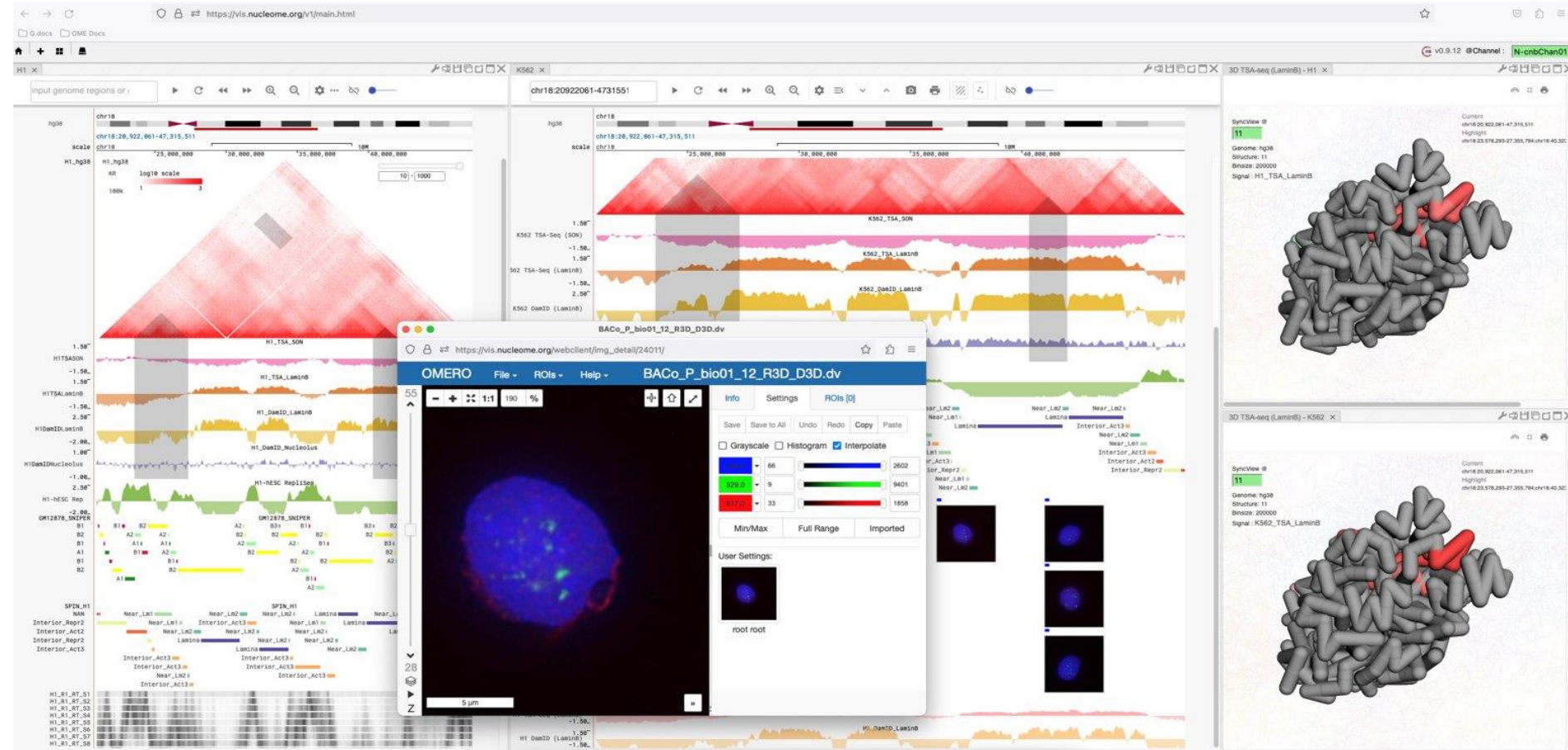
To filter rows you can use a query based on named columns. For example, to filter for rows where **shape** is greater than **6609251** add [?query=shape>6609251](#) to the URL.

For a more complex example, try [?query=\(shape>6609251\)&\(shape<6609254\)](#)

Table rows: 301.

Roi	shape	cell_id	fov	fov_amp_ind	x_um	y_um	z_um	region_size	hg38_chr	hg38_pos	umi_rol	purity_score	umi_seq	
2363501	6609251	1	1	2	6	15.2752294921875	19.9336328125	1.5	6	1	194631781	133144213232111000	0.619711344073634	TTGTATTAGTTTATGAATGCTGCCA
2363501	6609252	1	1	2	11	13.7585400390625	17.2252587890625	2.1	2	1	150295182	131323132424211000	0.833464665274033	AGGTAACAGCCGCCGGTGAAGCCCT
2363501	6609253	1	1	2	31	14.950224609375	17.116923828125	1.5	3	1	225100477	121224444112341000	1.41730804431887	CATTGAGAATGCTAGTTCGTATTCA
2363501	6609254	1	1	2	53	4.550068359375	12.133515625	1.8	3	1	159490134	134342412243342000	1.9908484714265902	AAGGTTGCATAGAGAAAGTTACCT
2363501	6609255	1	1	2	78	5.3084130859375	12.56685546875	1.8	3	1	23787304	132213213124423000	1.1026393293151202	GTATTAGTTTTGGTGTAGTATGCCA
2363501	6609256	1	1	2	232	13.650205078125	15.816904296875	1.5	4	1	153996406	241442114442434000	2.32473562719153	CTACGGATTATATGCTTCTTTCCCT
2363501	6609257	1	1	2	389	13.3252001953125	18.63361328125	2.4	4	1	173788354	224242441214242000	2.04123888148904	ATAGTTGAGGATCGAGCGGGGTAT
2363501	6609258	1	1	2	440	2.2750341796875	14.083544921875	1.5	2	1	201029547	233331413324421000	1.45461949740511	TGGACGATCGAGGGTACTGGTGCCA
2363501	6609259	1	1	2	478	4.116728515625	13.3252001953125	1.8	1	1	35436829	243342343234322000	2.10121851879784	CGTGC GTTTGATTAGGTTTTCCCT
2363501	6609260	1	1	2	673	4.3333984375	13.650205078125	1.5	4	1	27082504	333423423413433000	2.12571952800988	TTTAGGGTTGATGGTACTGTTGCCA
2363501	6609261	1	1	2	705	4.4417333984375	15.816904296875	2.1	2	1	92374532	312144232424141000	1.74120631396678	GTGTACTATTTTCACTCTGTGCA
2363501	6609262	1	1	2	780	5.3084130859375	13.866875	1.2	5	1	104368084	324221333321412000	2.53998829881878	TTGAGTAAAGTATCCGTCGTTGCCA
2363501	6609263	1	1	2	785	5.7417529296875	14.083544921875	1.8	6	1	52747838	333343312121144000	2.47414362805048	ATTCGTGCTTGTAGGTGCGGTCTCT
2363501	6609264	1	1	2	857	5.5250830078125	15.383564453125	1.8	2	1	178366318	333341322341142000	1.27093181216116	TAAGTAGGCCGTGAGACGTAGTCCT
2363501	6609265	1	1	2	993	5.0917431640625	13.216865234375	1.5	2	1	112059828	433212413342332000	1.05738785313283	TGCAATTTTGTAGTTGGTGGATGCCA
2363501	6609266	1	1	2	1081	3.3583837890625	11.916845703125	1.2	5	1	57737183	414422212422322000	0.958035105708203	AGCCATGCCCTTTTATTCATCCCT
2363501	6609267	1	1	2	1083	3.03337890625	14.950224609375	1.8	2	1	226248521	414234421211223000	1.33164034014242	GTTGAAAGTAGATTCGAAAAGTAT
2363502	6609268	1	1	2	35	13.7585400390625	6.6084326171875	1.8	2	2	11330392	124244422331142000	2.9277222459427406	ATTGTTTTAAAAGTCGAGTATTCCT
2363502	6609269	1	1	2	351	17.6585986328125	10.83349609375	1.2	4	2	154125297	214132314424122000	1.5961370210682602	GCTCAATACCAGTGCGGATGGTCA
2363502	6609270	1	1	2	613	10.183486328125	14.950224609375	1.8	1	2	79069289	222223234132422000	1.82441769480011	CTTGGAACCCAAATTTTACTTTCA

In Situ Sequencing - Linking Multimodal Metadata



Jhu et al, 2022, *Nature Methods*, 10.1038/s41592-022-01559-3
<http://vis.nucleome.org/entry/>



Imaging data for all studies published in IDR is available for download using the Aspera protocol (<https://idr.openmicroscopy.org/about/download.html>).



Datasets are published under the Creative Commons Attribution 4.0 International license (CC BY 4.0) or a more permissive license.



- **Share** — copy and redistribute the material in any medium or format.
- **Adapt** — remix, transform, and build upon the material for any purpose, even commercially.
- **Attribution** — credit must be given to the creator.



- [Java API](#) 
- [Python API](#)  
- [R API](#) 
- [CellProfiler & IDR](#) 
- [Fiji & IDR](#) 
- [Orbit & IDR](#) 
- [QuPath & IDR](#)
- [IDR notebooks](#)  

<https://github.com/ome/ABIS-GBI-imaging-course-07-2023> 

Validate chosen tool against data in IDR

IDR Studies Genes Phenotypes Cell Lines siRNAs Antibodies Compounds Organisms About

Public Public data

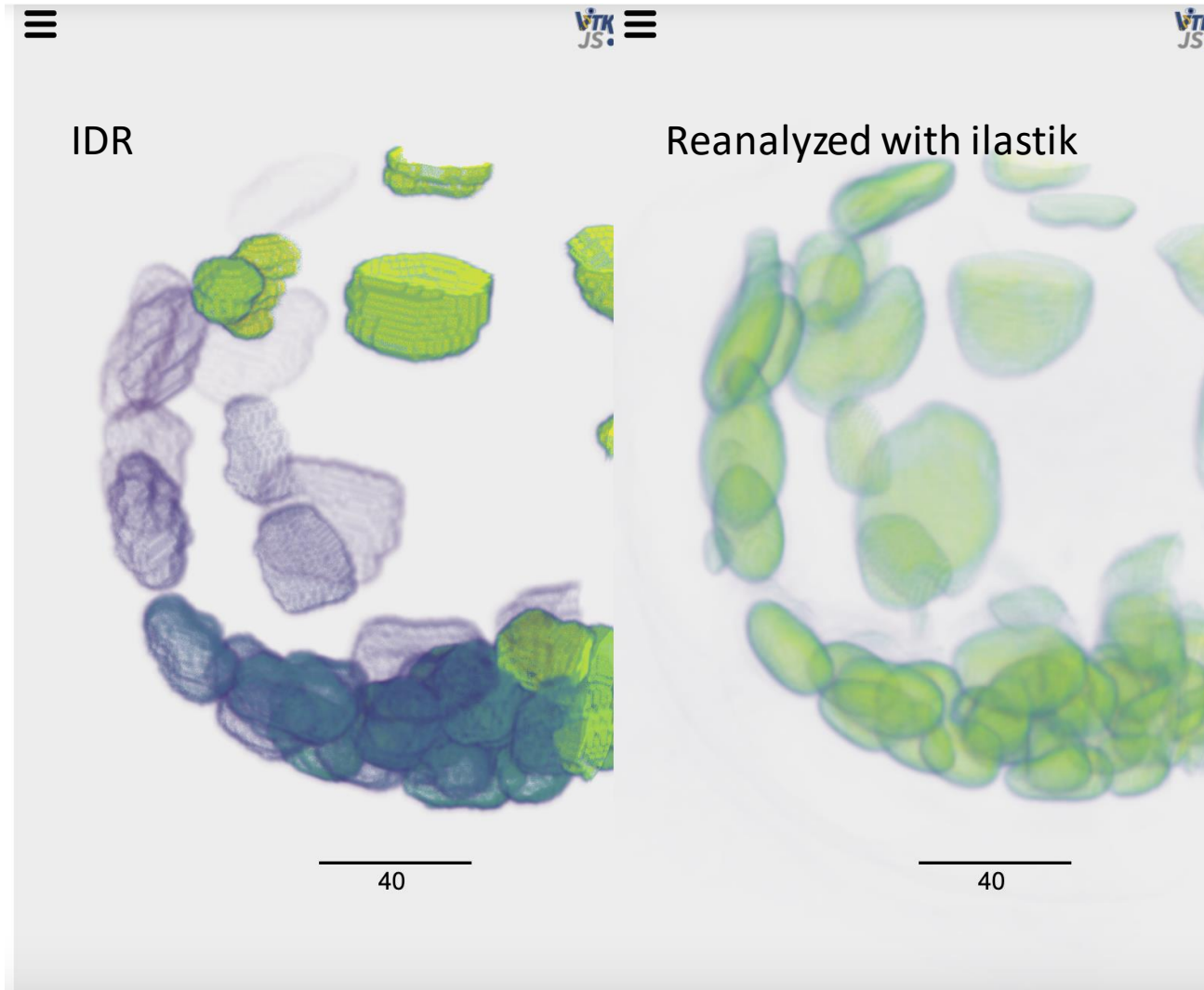
Explore Tags Shares

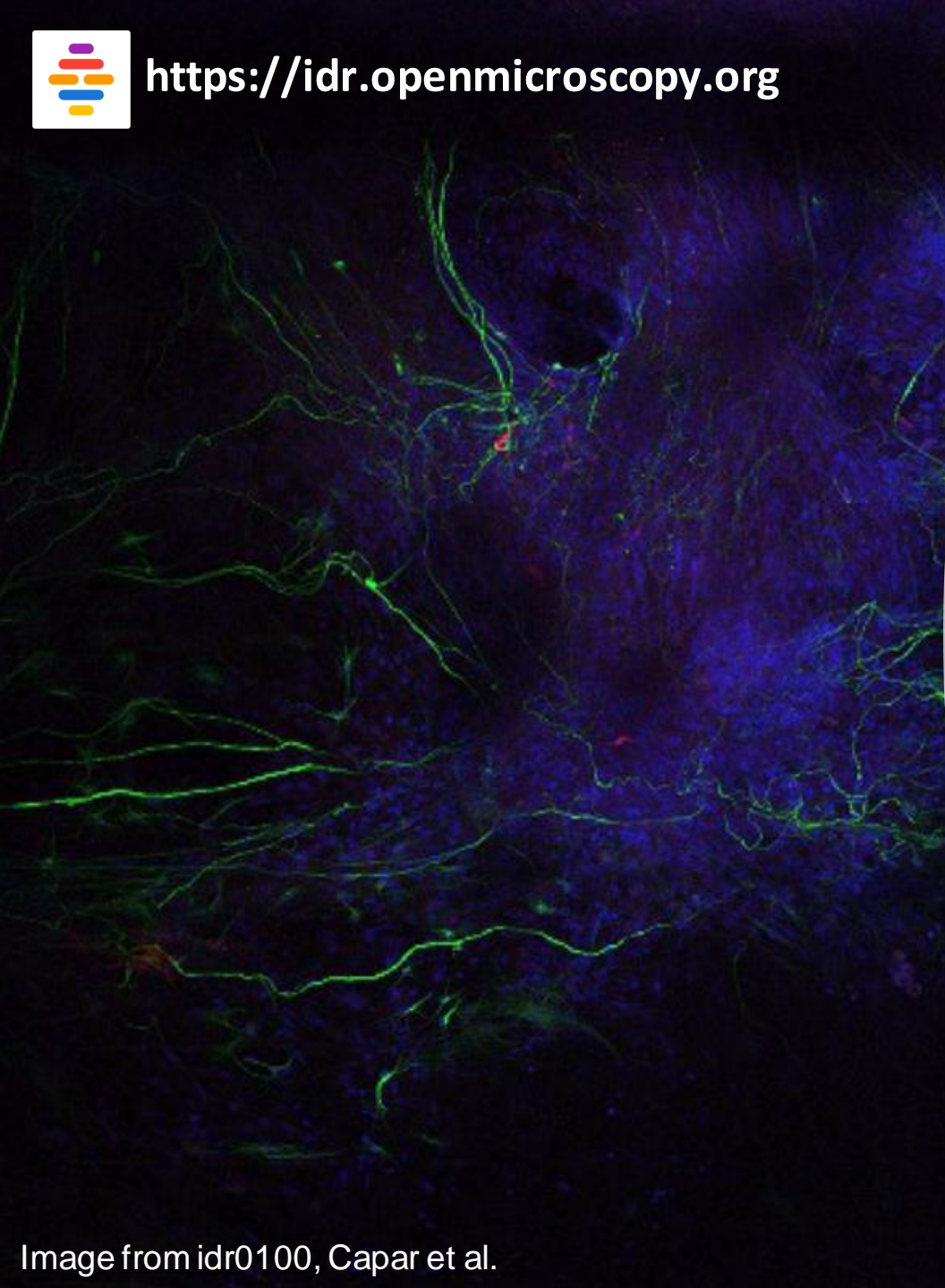
idr0062-blin-nuclearsegmentation/experimentA 5

- Acini 3
- Blastocysts 11
 - B1_C1.tif
 - B1_C2.tif
 - B2_C1.tif
 - B2_C2.tif
 - B3.tif
 - B4_C1.tif
 - B4_C2.tif
 - B4_C3.tif
 - B5_C1.tif
 - B5_C2.tif
 - B5_C3.tif
- E75 3
- E875 4
- Neural 1

Add filter

Compare ROIs in IDR to reanalysis output





Take-home message:

- IDR implements **FAIR** data
 - Findable
 - Accessible
 - Interoperable
 - Reusable
- IDR data can be used to **validate** tools

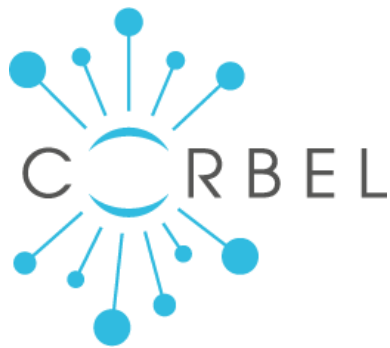
Thanks to the funders



Biotechnology and
Biological Sciences
Research Council



Chan
Zuckerberg
Initiative 



GLOBAL
BIOIMAGING
growing collaboration



Thanks to the IDR Team



**Jason
Swedlow**



**Frances
Wong**



**Dominik
Lindner**



**Josh
Moore**



**Jean-Marie
Burel**



**Petr
Walczysko**



**Will
Moore**



**David
Gault**



**Khaled
Mohamed**

Past members

**Simon Li
Mark Carroll
Riad Gozim
Eleanor Willams
Simone Leo
Gabry Rustici
Ola Tarkowska
Balaji Ramalingam
Rafael Carazo-Salas
Balint Antal
Anatole Chessel
Simon Jupp
Tony Burdett**



OME



**Sebastien
Besson**



**Alvis
Brazma**



**Ugis
Sarkans**



**Matthew
Hartley**