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

OME Team
University of Dundee
@openmicroscopy, @IDRNews
Welcome to IDR

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.

Search by:

- Name (IDR number)
- Type to filter values...

Most Recent (10)

- idr0094B Ellinger B et al.
  Identification of inhibitors of SARS-CoV-2 in-vitro cellular toxicity in human (Caco-2) cells using a large scale drug repurposing collection

- idr0094A Ellinger B et al.
  Identification of inhibitors of SARS-CoV-2 in-vitro cellular toxicity in human (Caco-2) cells using a large scale drug repurposing collection

- idr0091A Julou T et al.
  Subpopulations of sensorless bacteria drive fitness in fluctuating environments

- idr0089B Fischl H et al.
  Cold-induced chromatin condensation and nuclear retention of circadian mRNAs resets the circadian

https://idr.openmicroscopy.org
Welcome to IDR

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.

Most Recent (10)

- idr0094B Ellinger B et al.
  - Identification of inhibitors of SARS-CoV-2 in vitro cellular toxicity in human (Caco-2) cells using a large scale drug repurposing collection

- idr0094A Ellinger B et al.
  - Identification of inhibitors of SARS-CoV-2 in vitro cellular toxicity in human (Caco-2) cells using a large scale drug repurposing collection

- idr0091A Julou T et al.
  - Subpopulations of sensorless bacteria drive fitness in fluctuating environments

- idr0089B Fischl H et al.
  - Cold-induced chromatin condensation and nuclear retention of clock mRNAs resets the circadian clock

https://idr.openmicroscopy.org
Making image data FAIR

**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
## IDR metrics

### January 2022

<table>
<thead>
<tr>
<th>Category</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw data size</td>
<td>307 TB</td>
</tr>
<tr>
<td>Image files</td>
<td>30 M</td>
</tr>
<tr>
<td>5D Images</td>
<td>12 M</td>
</tr>
<tr>
<td>2D Planes</td>
<td>105 M</td>
</tr>
<tr>
<td>Organisms</td>
<td>15</td>
</tr>
<tr>
<td>Genes</td>
<td>89 K</td>
</tr>
<tr>
<td>Antibodies</td>
<td>11 K</td>
</tr>
<tr>
<td>Compounds</td>
<td>40 K</td>
</tr>
</tbody>
</table>
IDR studies and publications

>85 cross-published studies

cross-referenced via accessions and DOIs

https://idr.openmicroscopy.org/about/studies.html
Integrated studies

Thumbnails (of 5D Images)

Gene Product Targeting HCS

Genetic HCS

Geographic HCS

Chemical HCS

Histopathology

3D-Sim

Super-resolution

Experimental metadata

Biomolecular annotations

Analysis results

Cross-data browsing

Cloud analysis

Download (local analysis)
Linked Metadata

idr0012

**Attributes**

**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
<table>
<thead>
<tr>
<th>Ontologies</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Organism</td>
<td>NCBITaxon</td>
</tr>
<tr>
<td>Study Type</td>
<td>Experimental Factor Ontology (EFO)</td>
</tr>
<tr>
<td></td>
<td><a href="https://www.ebi.ac.uk/efo/">https://www.ebi.ac.uk/efo/</a></td>
</tr>
<tr>
<td>Screen Type (HCS)</td>
<td>Experimental Factor Ontology (EFO)</td>
</tr>
<tr>
<td></td>
<td><a href="https://www.ebi.ac.uk/efo/">https://www.ebi.ac.uk/efo/</a></td>
</tr>
<tr>
<td>Screen Technology Type (HCS)</td>
<td>Experimental Factor Ontology (EFO)</td>
</tr>
<tr>
<td></td>
<td><a href="https://www.ebi.ac.uk/efo/">https://www.ebi.ac.uk/efo/</a></td>
</tr>
<tr>
<td>Library Type (HCS)</td>
<td>Experimental Factor Ontology (EFO)</td>
</tr>
<tr>
<td></td>
<td><a href="https://www.ebi.ac.uk/efo/">https://www.ebi.ac.uk/efo/</a></td>
</tr>
<tr>
<td>Protocol</td>
<td>Experimental Factor Ontology (EFO)</td>
</tr>
<tr>
<td></td>
<td><a href="https://www.ebi.ac.uk/efo/">https://www.ebi.ac.uk/efo/</a></td>
</tr>
<tr>
<td>Imaging Method</td>
<td>Biological Imaging Methods Ontology (FBbi)</td>
</tr>
<tr>
<td></td>
<td><a href="https://www.ebi.ac.uk/ols/ontologies/fbbi">https://www.ebi.ac.uk/ols/ontologies/fbbi</a></td>
</tr>
<tr>
<td>Phenotype</td>
<td>Cellular Microscopy Phenotype Ontology (CMPO)</td>
</tr>
<tr>
<td></td>
<td><a href="https://www.ebi.ac.uk/cmpo/">https://www.ebi.ac.uk/cmpo/</a></td>
</tr>
<tr>
<td>Clinical/Pathology</td>
<td>SNOMED CT</td>
</tr>
<tr>
<td></td>
<td><a href="https://bioportal.bioontology.org/ontologies/SNOMED">https://bioportal.bioontology.org/ontologies/SNOMED</a> CT</td>
</tr>
<tr>
<td>Compounds</td>
<td>PubChem</td>
</tr>
</tbody>
</table>
Tissue Data: Human Protein Atlas

EMPIAR-10404

SARS-CoV-2 productively infects human gut enterocytes

Publication: SARS-CoV-2 productively infects human gut enterocytes
Science DOI: 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
Experimental metadata: Download xml

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
Open cloud and open analytics

In [4]:
session = create_http_session(idr_url.value)
organism = org_sel.value
idr_base_url = idr_url.value

[query_genes_dataframe, screen_to_phenotype_dictionary] = get_phenotypes_for_genelist(idr_base_url, session, go_gene_list)
display(HTML(query_genes_dataframe.to_html(escape=False)))

<table>
<thead>
<tr>
<th>Entrez</th>
<th>Ensembl</th>
<th>Key</th>
<th>Value</th>
<th>PhenotypeName</th>
<th>PhenotypeAccession</th>
<th>ScreenIds</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARPC2</td>
<td>[10109]</td>
<td>GeneName</td>
<td>[decreased cell numbers, geometric cell phenot...</td>
<td>[CMPO_0000274, CMPO_0000299, CMPO_0000276, CMPO_0000296, CMPO_0000297]</td>
<td>[206, 206, 206, 206, 206, 206, 206, 206, 206, 206]</td>
<td></td>
</tr>
<tr>
<td>ARPC1A</td>
<td>[10552]</td>
<td>GeneName</td>
<td>[strong decrease in rate of protein secretion]</td>
<td>[CMPO_0000261, CMPO_0000287, CMPO_0000288, CMPO_0000289]</td>
<td>[206, 206, 206, 206, 206, 206, 206, 206, 206, 206]</td>
<td></td>
</tr>
<tr>
<td>ARPC5L</td>
<td>[81873]</td>
<td>GeneName</td>
<td>[increased actin localised to the nucleus, e...</td>
<td>[CMPO_0000305, CMPO_0000348, CMPO_0000307, CMPO_0000308]</td>
<td>[1101, 1101, 1101, 1101, 1101, 1101, 1101, 1101, 1101]</td>
<td></td>
</tr>
<tr>
<td>ACTR3C</td>
<td>[653857]</td>
<td>EnsemblID</td>
<td>[cell death phenotype, mitosis delayed phenot...</td>
<td>[CMPO_0000023]</td>
<td>[1101]</td>
<td></td>
</tr>
<tr>
<td>ACTR3B</td>
<td>[57180]</td>
<td>GeneName</td>
<td>[binuclear cell phenotype]</td>
<td>[CMPO_0000213]</td>
<td>[1101]</td>
<td></td>
</tr>
<tr>
<td>HEL-68</td>
<td>[10552]</td>
<td>EnsemblID</td>
<td>[strong decrease in rate of protein secretion]</td>
<td>[CMPO_0000319]</td>
<td>[251]</td>
<td></td>
</tr>
<tr>
<td>ARPC4</td>
<td>[10093]</td>
<td>GeneName</td>
<td>[mild decrease in rate of protein secretion]</td>
<td>[CMPO_0000318]</td>
<td>[251]</td>
<td></td>
</tr>
<tr>
<td>ACTR3</td>
<td>[10096]</td>
<td>GeneName</td>
<td>[strong decrease in rate of protein secretion,...</td>
<td>[CMPO_0000319, CMPO_0000319]</td>
<td>[251, 803]</td>
<td></td>
</tr>
</tbody>
</table>

https://github.com/IDR/idr-notebooks
**Open cloud and open analytics**

[https://github.com/IDR/idr-notebooks](https://github.com/IDR/idr-notebooks) can be launched in **MyBinder**, just click the badge

![idr-notebooks](https://github.com/IDR/idr-notebooks)

- [https://github.com/ome/omero-guide-cellprofiler](https://github.com/ome/omero-guide-cellprofiler)
- [https://github.com/ome/omero-guide-python](https://github.com/ome/omero-guide-python)

...contain IDR-specific notebooks too!
Redesigning data access for re-analysis

From file download

... to scalable remote data access

Aspera

Zarr + napari

https://idr.openmicroscopy.org/about/download.html
Publicly available, curated studies... submitted by the community...
in a searchable, scalable platform...
that links metadata ...
and enables reanalysis ...
that can be deployed by others.
Thanks to the funders

UKRI Biotechnology and Biological Sciences Research Council

Chan Zuckerberg Initiative

Wellcome

EOSC-Life

CORBEL growing collaboration

GLOBAL BIOIMAGING

MULTI

EURO-BIOIMAGING
Thanks to the IDR Team

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

Jason Swedlow
Sebastien Besson
Dominik Lindner
Frances Wong

Josh Moore
Jean-Marie Burel
Petr Walczysko
Will Moore
David Gault

Alvis Brazma
Ugis Sarkans

OME

EMBL-EBI
Which diabetes related genes are expressed in the pancreas?

From Disease to Images

HUMANMINE
An integrated database of Homo sapiens genomic data

Which diabetes related genes are expressed in the pancreas?

IDR MULTIOMIC API

Images linked to gene PDX1


Developmental stage