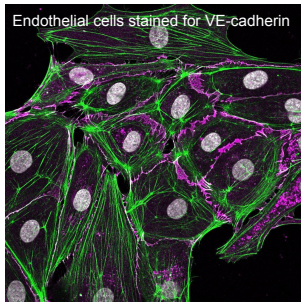
The Imaging Platform's data

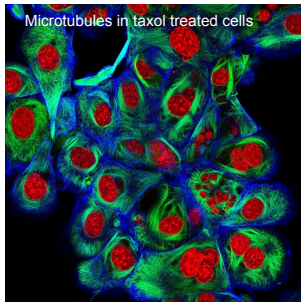
Alán F. Muñoz

2024/06/24

What kind of data can we extract from cells?

Why use cell images?





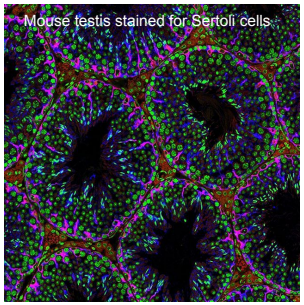
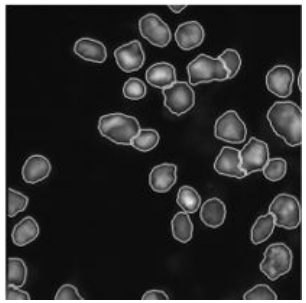
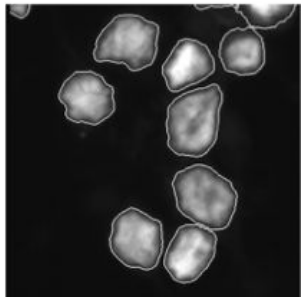


Image credit: @immunofluorescence (Derek Sung)

Treasure trove of information and offers single cell resolution





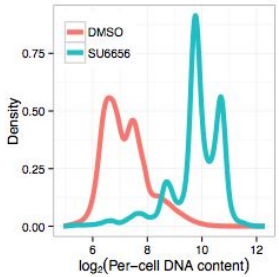


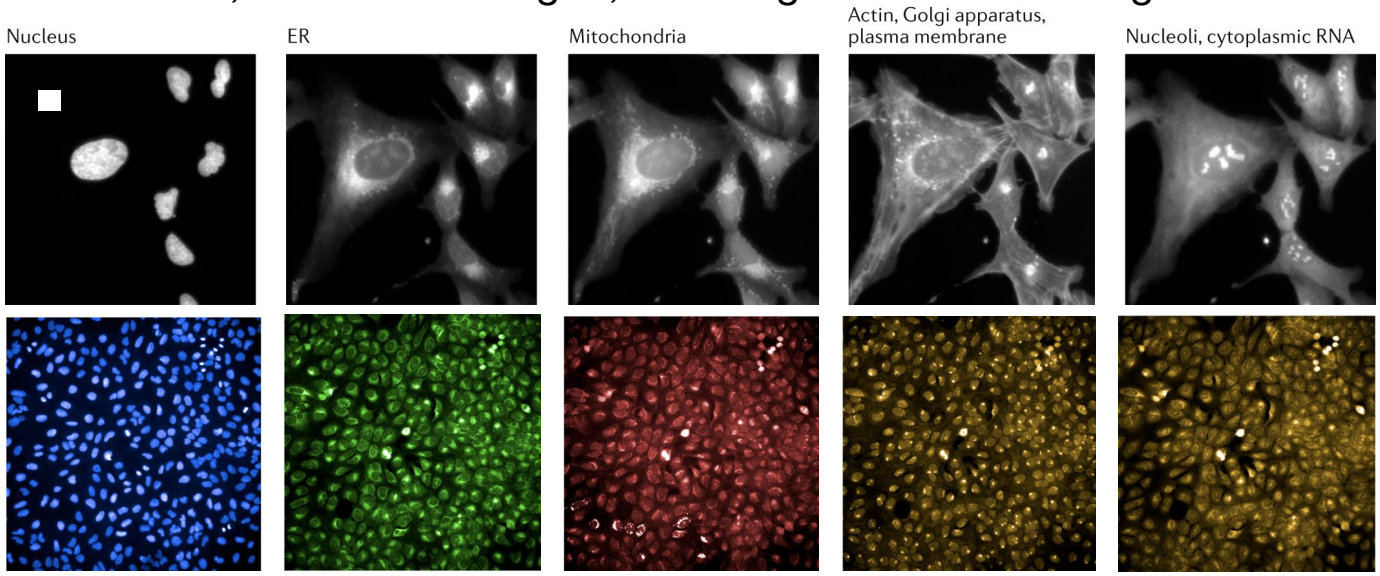
Image analysis is quantitative

Gandat. et al. (2019). Clin Cancer Res. CCR-19-1005
Wen, et al. (2012). Cell 150(3):575-89

Cell Painting Assay

Maximize information for morphological profiling

6 stains, 5 channels imaged, revealing 8 constituents/organelles:



Dyes	Hoechst 33342	Concanavalin A	Mitotracker Red	WGA/Phalloidin	SYTO14
Channels	Hoechst 33342	Alexa 488	Alexa 647	Alexa 568	488 Long

Gustafsdottir, et al. PLOS ONE 2013, Bray, et al. Nature Protocols 2016, Chandrasekaran et. al NRDD 2020

Cell Painting Gallery

The data from a wide variety of projects

- Raw images
- Illumination adjustments
- Morphological Profiles
    - Single-cell level
    - Aggregated profiles
- Segmentation masks
- Deep Learning Embeddings

Data open and available on AWS

Registry of Open Data on AWS

Cell Painting Gallery

bioinformatics | biology | cancer | cell biology | cell imaging | cell painting | chemical biology | computer vision | csv | deep learning | fluorescence imaging | genetic | high-throughput imaging | image processing | image-based profiling | imaging | machine learning | medicine | microscopy | organelle

Description

The Cell Painting Gallery is a collection of image datasets created using the Cell Painting assay. The images of cells are captured by microscopy imaging, and reveal the response of various labeled cell components to whatever treatments are tested, which can include genetic perturbations, chemicals or drugs, or different cell types. The datasets can be used for diverse applications in basic biology and pharmaceutical research, such as identifying disease-associated phenotypes, understanding disease mechanisms, and predicting a drug's activity, toxicity, or mechanism of action (Chandrasekaran et al 2020). This collection is maintained by the Carpenter–Singh lab and the Cimini lab at the Broad Institute. A human-friendly listing of datasets, instructions for accessing them, and other documentation is at the corresponding GitHub page about the Gallery.

Update Frequency

Typically when an associated publication is posted on biorxiv

License

CC0 1.0 Universal (CC0 1.0) Public Domain Dedication, but please do cite the corresponding publication for each dataset, as listed here.

Documentation

https://github.com/broadinstitute/cellpainting-gallery

Managed By

Carpenter-Singh and Cimini Labs at the Broad Institute

See all datasets managed by Carpenter-Singh and Cimini Labs at the Broad Institute.

Contact

cellpainting@broadinstitute.org

How to Cite

Resources on AWS

Description
Cell Painting data, comprising fluorescence microscopy cell images (TIFF), extracted features (CSV), and associated metadata (CSV and TXT).

Resource type
S3 Bucket

Amazon Resource Name (ARN)
`arn:aws:s3:::cellpainting-gallery`

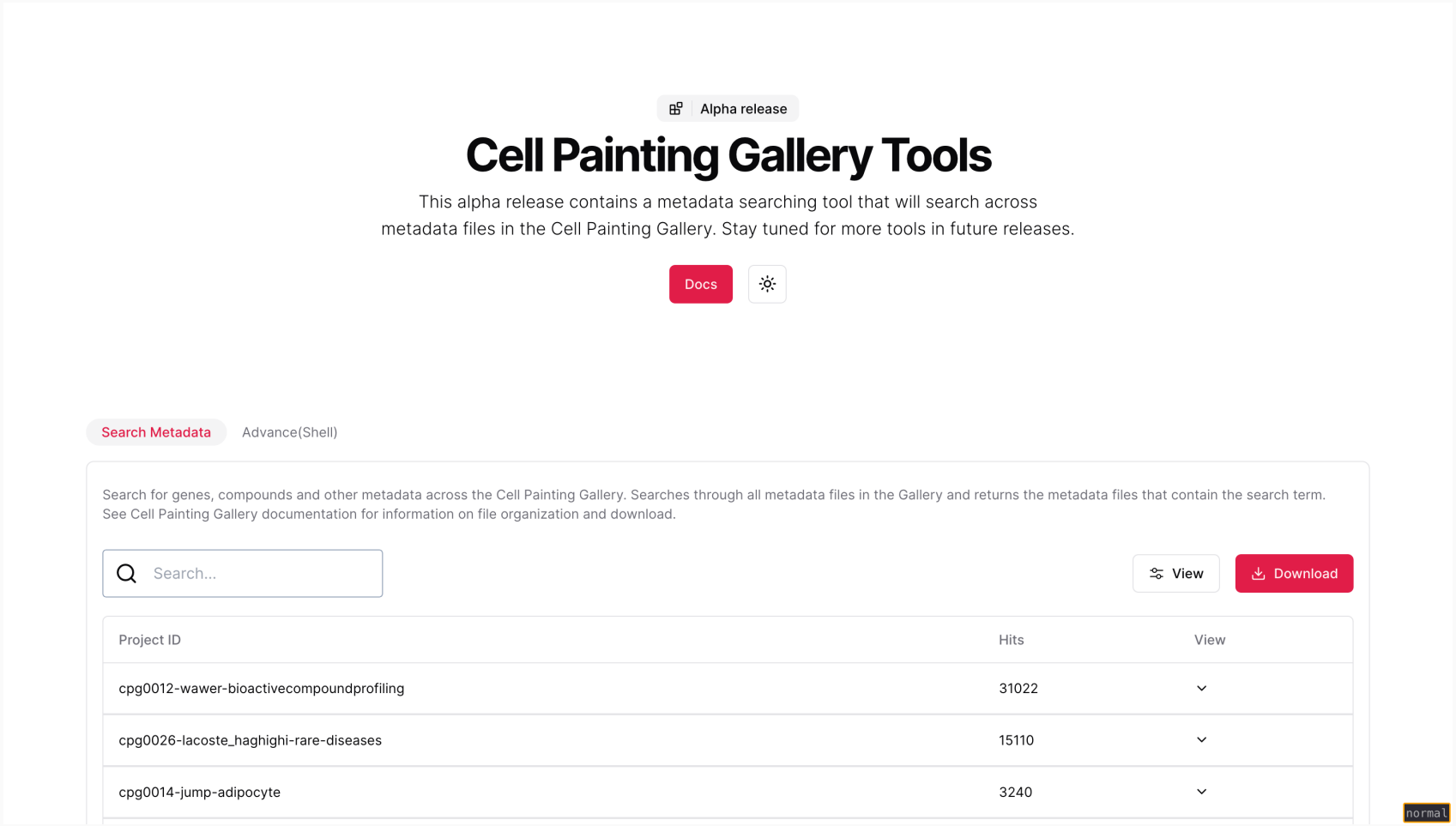
AWS Region
`us-east-1`

AWS CLI Access (No AWS account required)
`aws s3 ls --no-sign-request s3://cellpainting-gallery/`

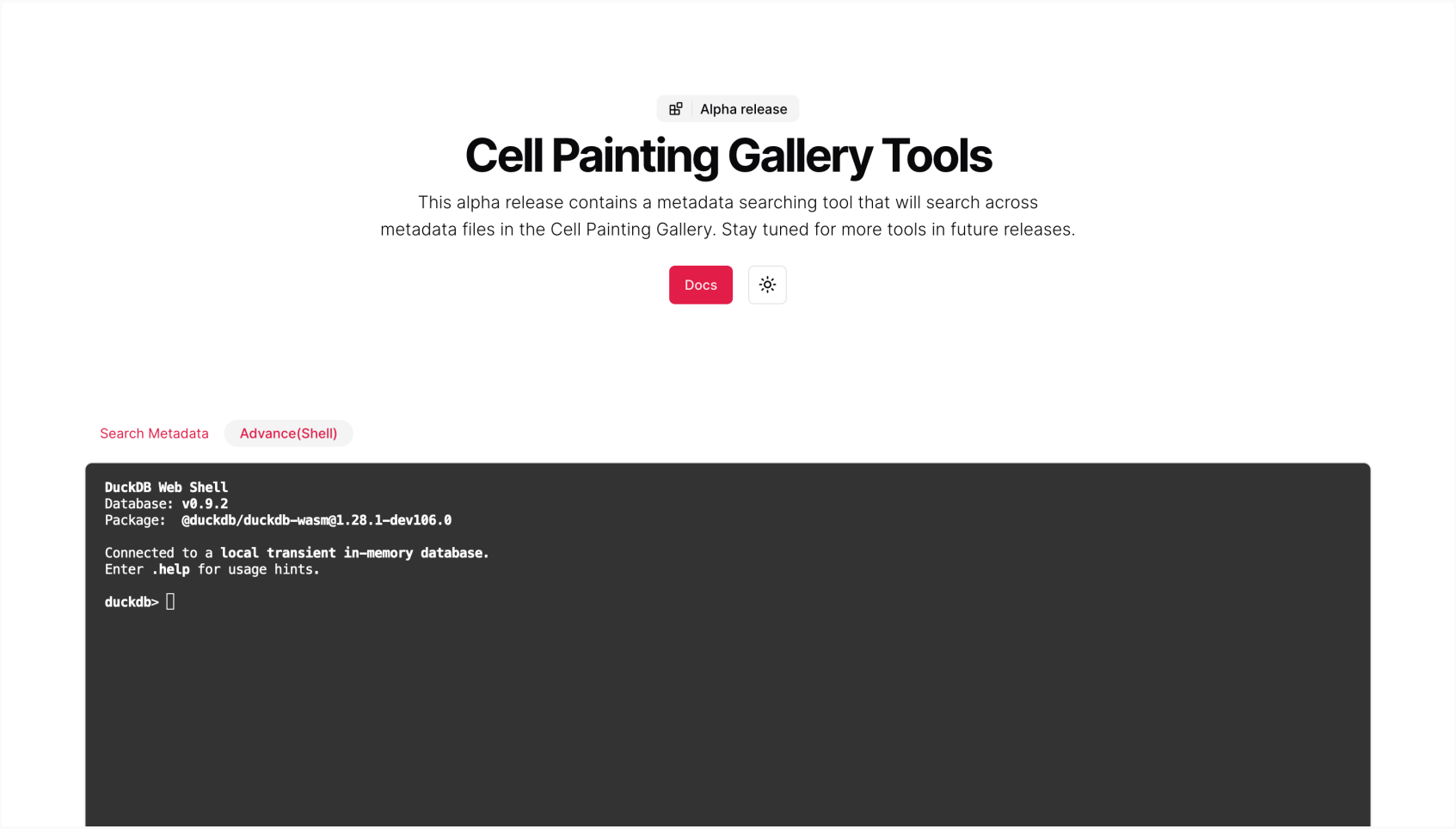
Explore
Documentation
Browse Bucket

normal

Search and browse metadata

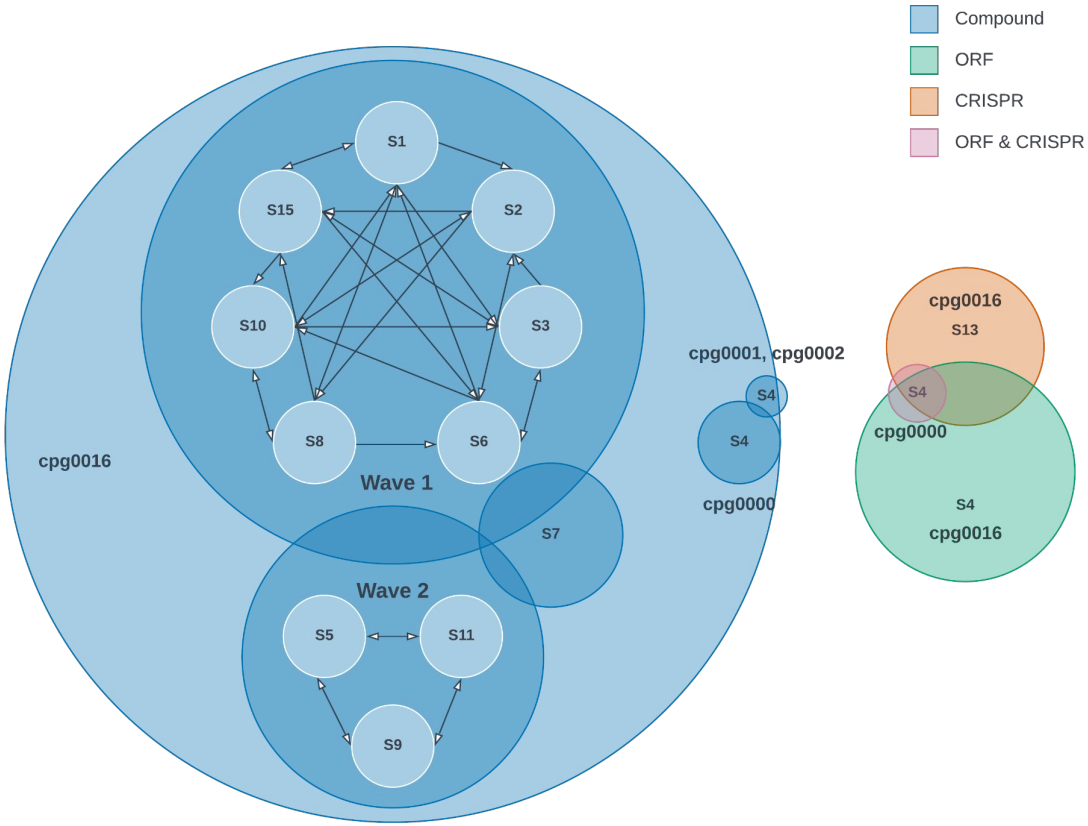


Includes in in-browser SQL



Joint-Undertaking for Morphological Profiling (JUMP)

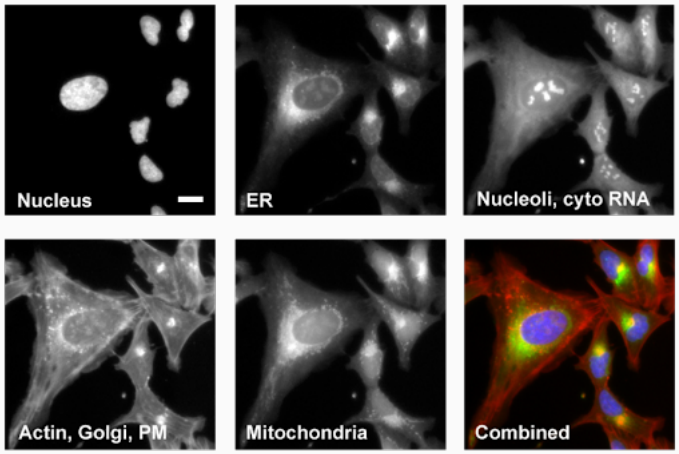
JUMP-CP dataset



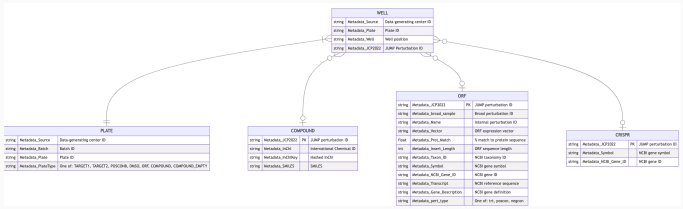
- Cell Painting images and profiles from ~1.5 billion cells
- ~140,000 perturbations
- 5 replicates
- 3 perturbation types: compound (85%), CRISPR knockout (5%) and ORF overexpression (10%)
- Data generated across 12 sites using 5 different microscopes.
- More details: Chandrasekaran et al. (2023) biorxiv

There is plenty of information to explore

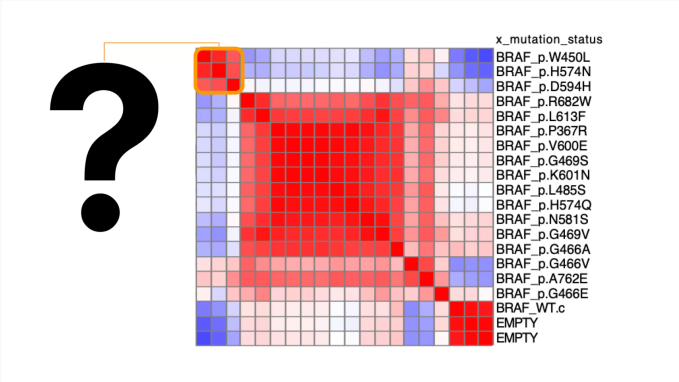
- Raw images



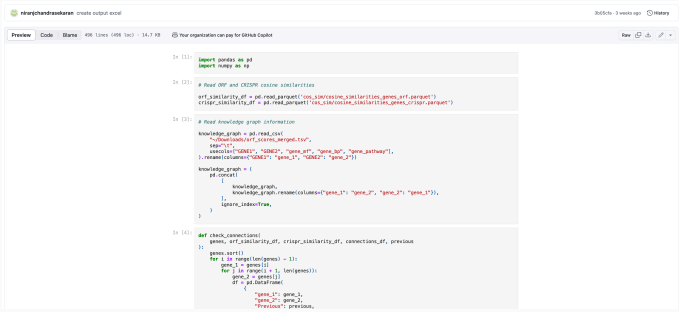
- Metadata tables



- Morphological profiles



- Notebooks



Similar morphological profiles

home / data

content

50 rows where Gene/Compound = "MYT1" sorted by Similarity descending

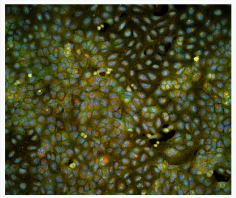
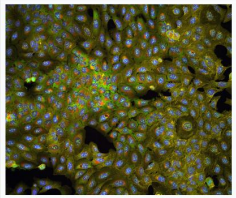
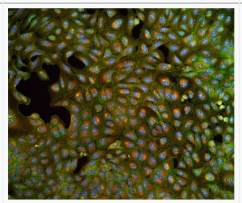
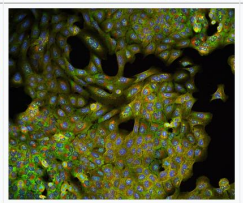
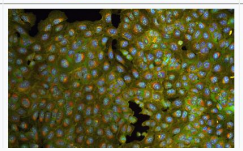
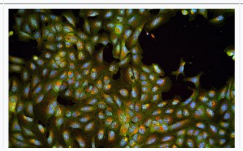
Gene/Compound = MYT1

- column - =

Apply

View and edit SQL

This data as json, CSV (advanced)

Link	rowid	index	Gene/Compound	Match	Gene/Compound Example	Match Example	Similarity ▲	Match resources	JCP2022
47551	47551	47550	MYT1	SOS1			0.7195221781730652	External	JCP2022_804400
47552	47552	47551	MYT1	IL13RA2			0.6855948567390442	External	JCP2022_804400
47553	47553	47552	MYT1	NLRP12			0.678930401802063	External	JCP2022_804400

normal

Most distinctive features for a given gene

67,700 rows

- column - = Apply

✎ View and edit SQL

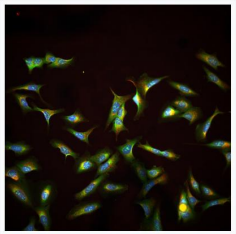
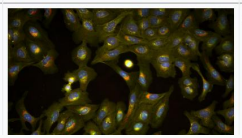
This data as json, CSV (advanced)

Channel 11 ✖

- RNA 12,350
- Mito 12,050
- ER 11,750
- DNA 11,450
- AGP 11,150
- - 5,100
- OrigDNA 800
- OrigRNA 800
- OrigAGP 750
- OrigER 750
- OrigMito 750

Mask 4 ✖

- Nuclei 19,100
- Cells 17,500
- Cytoplasm 17,050
- Image 14,050

Link	rowid ▼	index	Mask	Feature	Channel	Statistic	Gene/Compound	Metadata_image	Median	JCP2022
1	1	0	Cells	Correlation_K_Mito	AGP	8.349928508374835e-29	AVPR2		-0.21309228714749423	JCP2022_905491
2	2	1	Cells	Correlation_K_Mito	AGP	3.325815194945794e-19	EEF1AKMT2		0.13034929387328936	JCP2022_912740

normal

Image exploration

12 rows where Gene/Compound in ["MYT1", "PROZ"] sorted by rowid

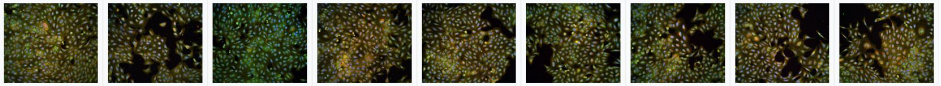
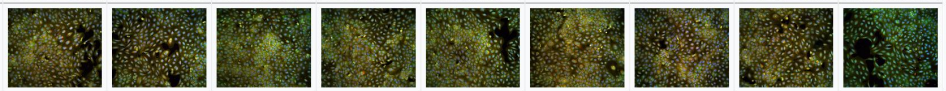
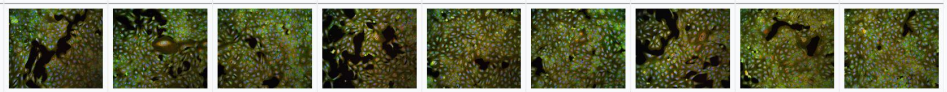
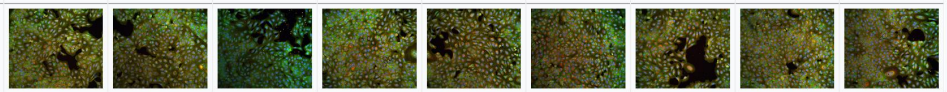
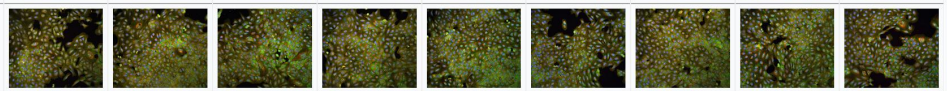
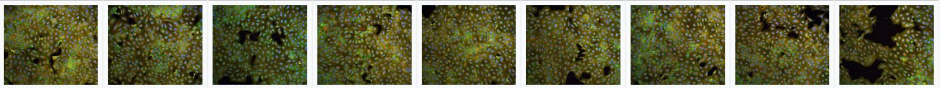
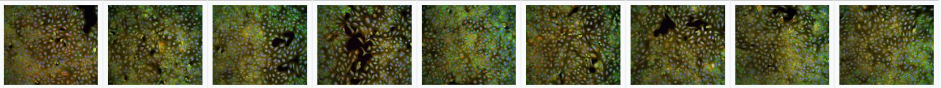
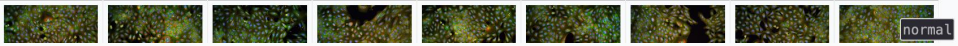
Gene/Compound | in | MYT1, PROZ

- column - | = |

Apply

✎ View and edit SQL

This data as json, CSV (advanced)

Link	rowid ▼	index	Gene/Compound	External resources	Metadata_JCP2022	Foci 0	Foci 1	Foci 2	Foci 3	Foci 4	Foci 5	Foci 6	Foci 7	Foci 8
6	6	5	PROZ	External	JCP2022_805564									
1055	1055	1054	MYT1	External	JCP2022_804400									
9671	9671	9670	PROZ	External	JCP2022_805564									
10704	10704	10703	MYT1	External	JCP2022_804400									
19344	19344	19343	PROZ	External	JCP2022_805564									
20390	20390	20389	MYT1	External	JCP2022_804400									
29042	29042	29041	MYT1	External	JCP2022_804400									
29521	29521	29520	PROZ	External	JCP2022_805564									

Methods of access: Software engineers

We provide Python packages to integrate in your analyses

- **`broad_babel:`** Synonyms and treatment metadata
- **`jump_portrait:`** Download images from gene names/chemical SMILES or fetch into your notebooks

Resources

Resources

- **`broad.io/jump:`** JUMP tutorials and info hub
- **`broad.io/monorepo:`** Imaging Platform Computational tools
- **`broad.io/cpg:`** For metadata exploration website
- **`slides:`** github.com/afermg/2024_06_broadhacks