



Populating EOSC-Life: Success stories from the
Demonstrator Projects

Reusing published microscopy images to study nucleolus biology

Beatriz Serrano-Solano University of Freiburg (former EMBL)

Jean-Karim Hériché EMBL

Yi Sun EMBL

Jean-Marie Burel University of Dundee

Björn Grüning University of Freiburg



Your presentation

The presentation should be 20 min long with an additional 5 min for questions.

Please remember to keep the content accessible to a general audience

In addition to describing the project, please make sure to also cover the following points (as suggested by the following slides):

- Impact of your project on the EOSC
- Sustainability of your project and future work
- Your interactions with the EOSC-Life technical experts
- Your experience with EOSC-Life





What is the biological background?

Goal

Looking at phenotypes to find potential regulators of nucleolus biology

Gene silencing

[2.10 TB] **Chromosome condensation**
[0.08 TB] **DNA damage**
[0.38 TB] **CellMorph**
[0.65 TB] **YAP activation**
[3.04 TB] **Secretion**
[13.90 TB] **MitoCheck**



Overexpression

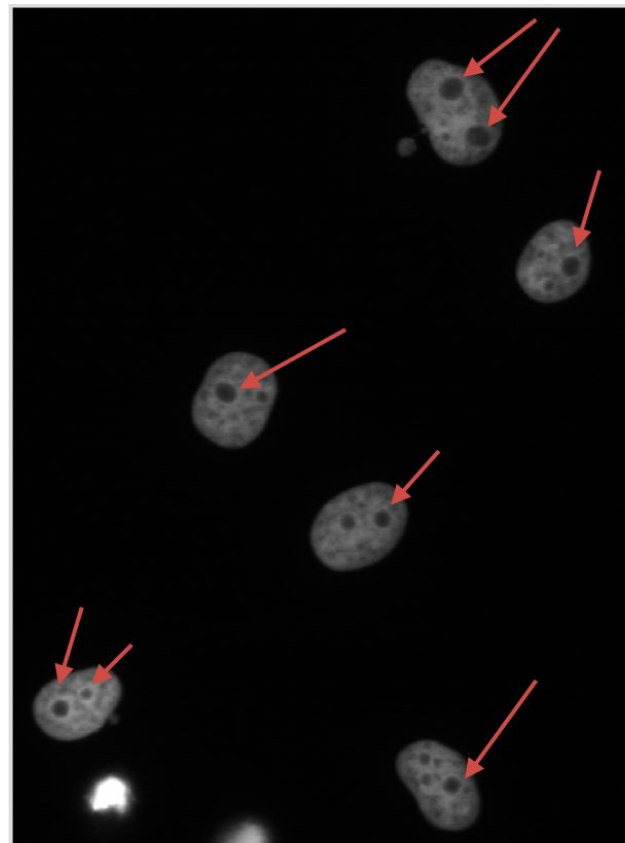
[1.40 TB] **Nuclear bodies**
[0.49 TB] **Signalling pathways**

Compound

[3.19 TB] **Bioactive compound**
[0.10 TB] **Drug response**

Interests

- number of nucleoli/nucleus
- shape
- spatial distribution



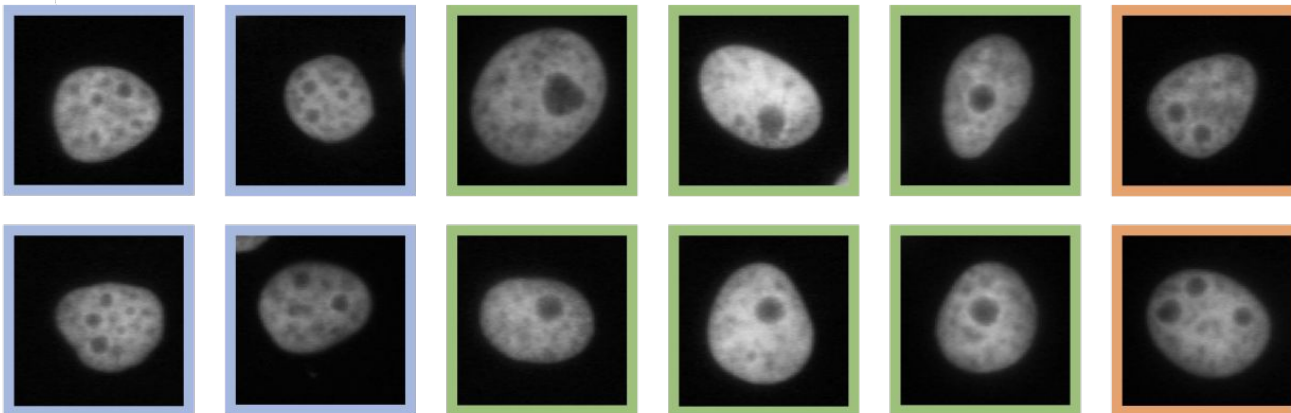
DNA channel





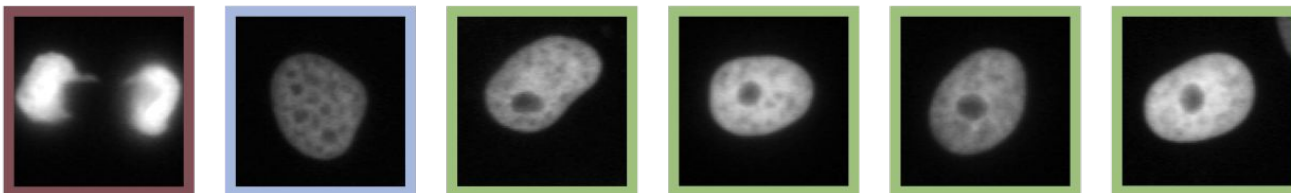
How do the possible phenotypes look like?

Several nucleoli evenly distributed in space



Nucleoli distributed close to the nuclear membrane

Absence of nucleolus



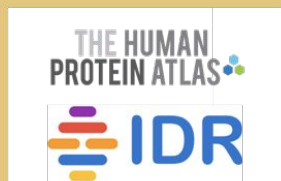
Large nucleolus



How does the image analysis pipeline look like?



Data Access



Segmentation & Feature Extraction



Downstream Biological Analysis



IDR Download

19 Galaxy tools
integrating 22 CellProfiler modules

INTERACTIVE NOTEBOOK

AUTOMATED WORKFLOW



What are the challenges?



image.sc

Community Partners


Metadata standards discussion at the IDR

Data Management metadata, idr

jkh1 2 Sep '19

For those of you interested in metadata standardization, a discussion has been initiated at the IDR over here:

github.com/idr-contrib/community 8

 **Issue: Metadata standards** 8

opened by beatrizserrano on 2019-09-17 8

I'd like to point out some issues that I've found related to the lack of metadata standards:
The number of columns is...

5 Open ✓ 23 Closed Author Label

Download several images as a single file

4 May

ero

ec 2019

s

9

30 Mar

r0043

9 Jul 2019

een idr0006

28 Apr

OMERO.gateway

1 Aug 2019

ared images with EBImage

1 Aug 2019

ded

018

and image annotation?

#14 by Running-z was closed on 13 Jul 2018.



IDR tool to access images from Galaxy



- Triggered at the **WP1 Hackathon in Berlin** (Nov 2019)
- Developed at the **Hackathon in Freiburg** (Feb 2020)



IDR/OMERO Download - download images from any OMERO instance using image IDs (Galaxy Version 0.43)

★ Added Versions Options

Local OMERO instances accessible

Which OMERO instance to connect?
IDR

By default, the tool will download a tarball containing individual images from IDR into your Galaxy history. If you need to connect to your own instance, set your connection username and password from User->Preference->Manage Information

How would you like to specify the IDs of images to download?
As text (comma-separated list of IDs or a valid IDR link)

Image IDs to download

You can enter a single image-id, or a comma (or '|')-separated list of IDs. Alternatively, you can paste here a link to an image selection obtained through the IDR webclient.

Name of the channel to download

For all image IDs only the specified channel will be downloaded. If left empty, the first channel (whatever this is) will be downloaded by default.

z-plane of images to download
0

Image frame to download
0

Limit the download to a selected region of the image?
No, download the entire image plane

Skip failed retrievals?
Yes No

By default the tool will fail with an error on the first non-retrievable image ID. Here, you can choose to skip non-retrievable image IDs and continue downloading the available ones instead. The error log will contain warnings about failed IDs in this case.

Download images in a tarball?
Yes No

Define ROI



CellProfiler integrated into Galaxy



Galaxy / Europe Analyze Data Workflow Visualize Shared Data Help User Using 1%

Tools ☆ ↑

search tools ✕

image

GeneSeqToFamily preparation
converts data for the workflow

Image Converter

Landmark Registration Landmark Registration

Adapt an elastic transformation to a new image size with bUnwarpJ

Apply elastic transformation with bUnwarpJ

Operate on pixels with a mathematical expression

Adjust threshold of binary image

Compose two elastic transformations into a raw transformation with bUnwarpJ

Sharpen

Convert to binary (black and white)


Convert elastic transformation to raw with bUnwarpJ 👆

Compare two raw deformations by warping index with bUnwarpJ

COVID-19 research!

Want to learn the best practices for the analysis of SARS-CoV-2 data using Galaxy? Visit the Galaxy SARS-CoV-2 portal at covid19.galaxyproject.org. We mirror all public SARS-CoV-2 data from ENA in a Galaxy data library for your convenience. The Galaxy community also created COVID-19 related trainings and we also maintain a running document with recent news.

James Taylor (1979-2020) believed that scientific progress can best be sustained through the mentoring of students and junior faculty.



Design by [Rebekka Paisner](#)

To ensure implementation of this vision, the Galaxy community has established a foundation—Junior Training and Educational Connections Hotspot (JTech). JTech's mission is to (1) assist graduate students to participate in computational biology and data science conferences, and (2) organize and host mentoring sessions between senior and junior faculty members at high-profile meetings.

To make this happen we are accepting contributions. More details can be found on [the @jtx page in the Galaxy Hub](#). Please, help us continue what James has started.

Donate now

"Anyone, anywhere in the world should have free, unhindered access to not just my research, but to the research of every great and enquiring mind across the spectrum of human understanding." – Prof. Stephen Hawking

History ↺ + 📄 ⚙️

search datasets ✕

Unnamed history

(empty) 🗨️

i This history is empty. You can load your own data or get data from an external source

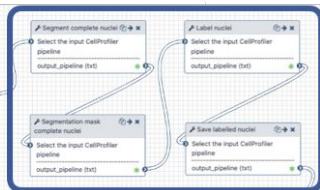
News OPEN CHAT >



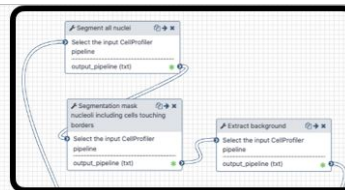
How does the automated part of the workflow look like?



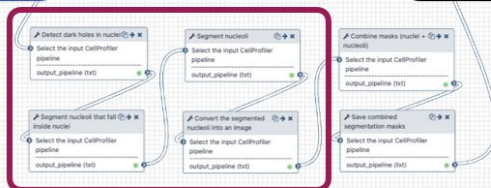
NUCLEI



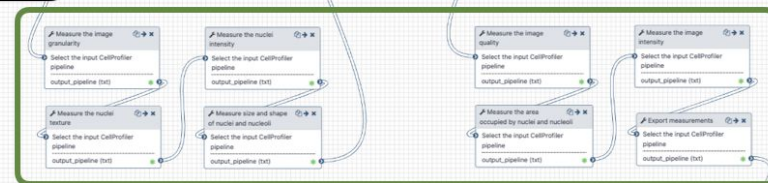
BACKGROUND



NUCLEOLI



FEATURE EXTRACTION



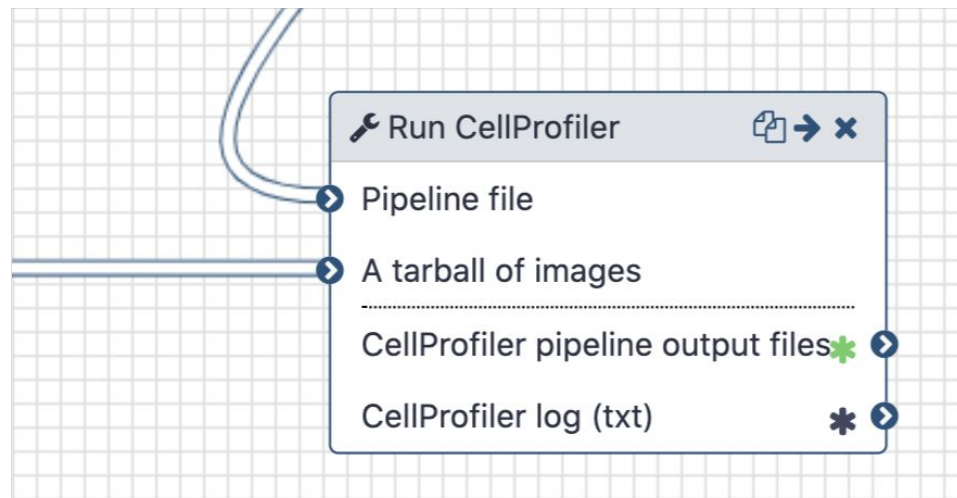
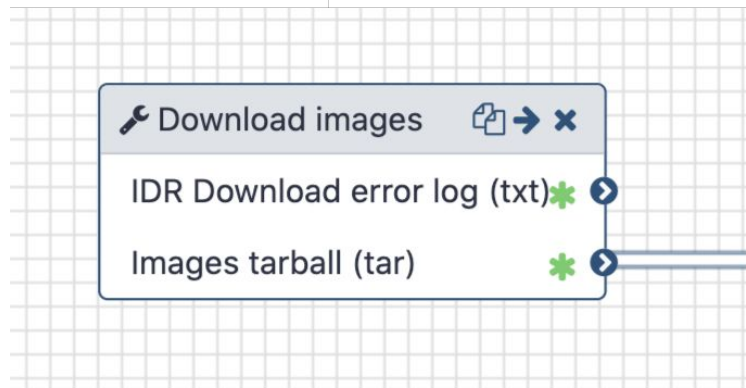
Starting modules

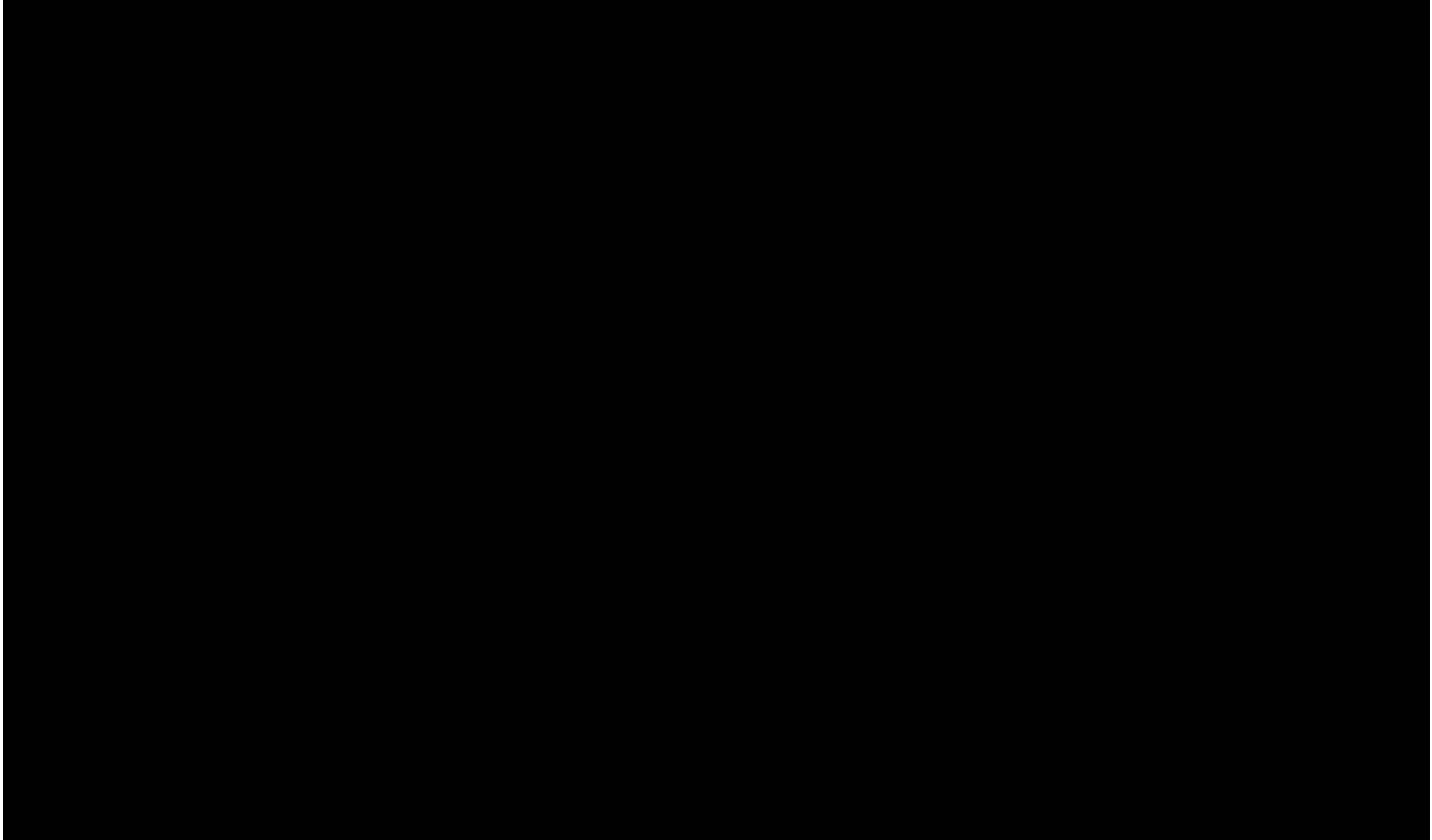
Download images
IDR Download error log
Images toolbar

Run CellProfiler
Pipeline file
A tarball of images
CellProfiler pipeline output files
CellProfiler log

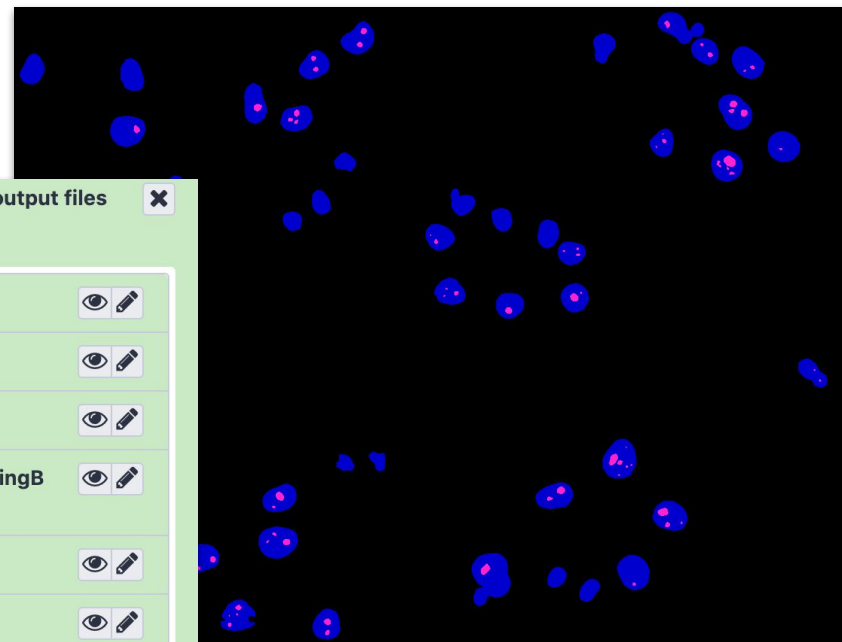
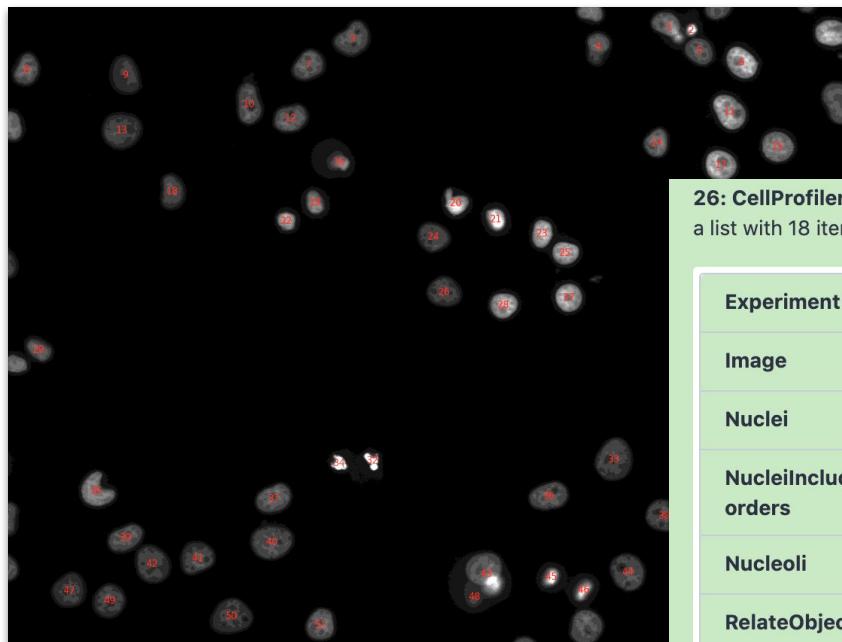


What are the inputs and outputs of the workflow?





















What is the output of the workflow?



26: CellProfiler pipeline output files ✕
a list with 18 items

Experiment	 
Image	 
Nuclei	 
NucleiIncludingTouchingB orders	 
Nucleoli	 
RelateObjects	 
plate2_2_007_[Well_27,_Fie ld_1_(Spot_27)]_295918_ 0_0_1344_1024_combinedMask	 
plate2_2_007_[Well_27,_Fie	 



How can anyone use the workflow?



Workflow Hub

Workflow Hub

Home / Workflows Index / Image analysis using CellProfiler (EOSC-Life D6)

Image analysis using CellProfiler (EOSC-Life D6) Version 1

Request Contact Unsub

Stable

This workflow has been created as part of Demonstrator 6 of EOSC-Life WP3 and is focused on reusing publicly available RNAi screens to gain insights into the nucleolus biology. The workflow downloads images from the Image Data Resource (IDR), performs object segmentation (of nuclei and nucleoli) and feature extraction of the images and objects identified.

NUCLEI

NUCLEOLI

BACKGROUND

FEATURE EXTRACTION



Tutorial




Image analysis using Galaxy

Requirements

Before diving into this topic, we recommend you to have a look at:

- [Introduction to Galaxy Analyses](#)

Material

Lesson	Slides	Hands-on	Input dataset	Workflows	Galaxy tour	Galaxy instances
Analyse HeLa fluorescence siRNA screen						
Introduction to image analysis using Galaxy						
Nucleoli segmentation and feature extraction using CellProfiler						

Nucleoli segmentation and feature extraction using CellProfiler

By: Beatriz Serrano-Solano, Jean-Karim Hériché

Overview

Questions

- How do I run an image analysis pipeline on public data using CellProfiler?
- How do I analyse the DNA channel of fluorescence siRNA screens?
- How do I download public image data into my history?
- How do I segment and label cell nuclei?
- How do I segment nucleoli (as the absence of DNA)?
- How do I combine nuclei and nucleoli into one segmentation mask?
- How do I extract the background of an image?
- How do I relate the nucleoli to their parent nucleus?
- How do I measure the image and object features?
- How do I measure the image quality?

Objectives

- How to download images from a public image repository.
- How to segment cell nuclei using CellProfiler in Galaxy.
- How to segment cell nucleoli using CellProfiler in Galaxy.
- How to extract features for images, nuclei and nucleoli.

Requirements

- [Introduction to Galaxy Analyses](#)

Time estimation: 4 hours

Supporting Materials

Slides Workflows Available on these Galaxies

Last modification: Dec 27, 2020

Introduction

Get data

Start CellProfiler pipeline

Object segmentation

Background extraction

Feature extraction

Run the CellProfiler pipeline

Conclusion

References

Feedback

Citing this Tutorial

Introduction

The nucleolus is a prominent structure of the nucleus of eukaryotic cells and is involved in ribosome biogenesis and cell cycle regulation. In DNA staining of cells, nucleoli can be identified as the absence of DNA in nuclei (Fig. 1). Phenotypes caused by reduced gene function are widely used to elucidate gene function and image-based RNA interference (RNAi) screens are routinely used to find and characterize genes involved in a particular biological process. While screens typically focus on one biological process of interest, the molecular markers used can also inform on other processes. Re-using published screens image data can then be a cost-effective alternative to performing new experiments.

In particular, regardless of the targeted biological process, many screens include a DNA label and therefore can also reveal the effect of gene knock-downs on nucleoli.





Impact of the project on EOSC

- **Exemplary EuBI workflow** in a cross-RI environment (Galaxy)
 - **CellProfiler** available in Galaxy to **all communities**
 - **IDR/OMERO tool** available in Galaxy
 - Opening access to data to **all communities**
- **Workflow** published in the Workflow Hub
 - Reusability
 - Adapted for different inputs / biological problems
- **Training material** available in the Galaxy Training Network





Future work/Sustainability of the project outcome

- Future work
 - Workflow **testing**
 - Scientific **publication** in progress
- Sustainability
 - Large **Open Source communities** (Bioconda, Galaxy)
 - Training material available to kickstart an **imaging community around Galaxy**





Interaction with EOSC-Life technical experts

Jean-Karim Hériché	WP2	Supervisor of the project
Yi Sun	WP2	Integration of tools into Galaxy
Jean-Marie Burel	WP1	WP1 Hackathon Berlin, Hackathon Freiburg, regular discussions on chat, GitHub, calls...
Björn Grüning	WP2, WP7	





Your experience with EOSC-Life

- **Support** from experts in different fields
 - Close feedback loop/physical proximity and hackathons
- **Access to cloud** resources because of EOSC-Life
- **Connect** Galaxy and imaging communities
 - Bringing Galaxy to the imaging community
 - Galaxy makes imaging tools accessible to other communities





Acknowledgements

Jean-Karim Hériché

EuBI

Yi Sun

EuBI

Jean-Marie Burel, IDR team

EuBI

Björn Grüning, Freiburg Galaxy Team

ELIXIR

