Programme of the day

- General introduction to OMERO
- **Coffee**
- Organizing Image Data
- **Lunch**
- Working with Image Data
- **Coffee**
- ImageJ
Talk Outline

- The Problem
- Open-source software tools
  - Image data access: OME Files & Bio-Formats
  - Image data management & sharing: OMERO
- Data Sharing, Provenance, and Integration
  - Scalability examples
  - Image Data Repository (IDR)
- Practical Use case
THE PROBLEM
The Image Problem... is Ubiquitous

- Organelles
- Cells
- Dynamics
- Physiology

- Lead Discovery
- Pathology
- In Vivo

A pretty picture?
A measurement?
A resource?
OPEN SOURCE TOOLS
What We Do

OME DATA MODEL

OME-XML
OME-TIFF
OME FILES

Open File Formats

BIO-FORMATS

Proprietary File Format Translation

OMERO

Image Data Management
BIO-FORMATS: Proprietary File Conversion

Raw Data

Processed Data

BIO-FORMATS

600,000+ FILES

30,000+ DATASETS

5.7+ TERABYTES

Apps using Bio-Formats were started >100,000 times per day in 2015
During your PhD and beyond…

- Store and organize original data
- Present the data to your supervisor and others
- Prepare poster and presentation
- Publish the data
- Cooperate on the data
- Analyze data with different tools
- Leave the data in a manageable state even
- Re-use knowledge
The “Scientific Data” Paradigm

Gray et al, 2005, Scientific Data Management in the Coming Decade, Microsoft Research
The Extensible OMERO Platform

- Acquisition
- Clients
- Processing

INTERNET

SERVER

OMERO

- Image Data
- Search
- Analytics
- Metadata

Plugins Welcome
Examples of Analysis Integration

- FLIMfit – fluorescence lifetime fitting (Matlab)
- WND-CHRM -- weighted nearest neighbor machine learning (Python)
- ThunderSTORM and PALMSiever – Localisation SRM (ImageJ, Matlab)
- OMERO2CV – LSFM Multi-View Reconstruction (C++, OpenCV, ITK)
- uTrack – Globally optimised object tracking (Matlab)
- CellProfiler – HCS segmentation and features (Python)
- mTools – Otsu, basic segmentation (Matlab)
- **ImageJ/Fiji, Icy** – Pluggable, desktop Image processing tools (Java)
- Columbus Acapella® -- commercial Big Data processing…

→ See Webinars at glencoesoftware.com for Matlab, Jupyter examples
DATA SHARING, PROVENANCE, INTEGRATION…
# Security Model

<table>
<thead>
<tr>
<th>PERMISSIONS</th>
<th>Read</th>
<th>Annotate</th>
<th>Write</th>
<th>Privacy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Private</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>You</td>
</tr>
<tr>
<td>Group-Read</td>
<td>✔</td>
<td>✗</td>
<td>✗</td>
<td>You</td>
</tr>
<tr>
<td>Group-Annotate</td>
<td>✔</td>
<td>✔</td>
<td>✗</td>
<td>Your Group</td>
</tr>
<tr>
<td>Group-Write</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>Your Group</td>
</tr>
<tr>
<td>Public-Read</td>
<td>✔</td>
<td>✗</td>
<td>✗</td>
<td>Anyone</td>
</tr>
<tr>
<td>Public-Annotate</td>
<td>✔</td>
<td>✔</td>
<td>✗</td>
<td>Anyone</td>
</tr>
<tr>
<td>Public-Write</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>Anyone</td>
</tr>
</tbody>
</table>
## Tang Mitosis/Apoptosis ver. II - Dataset (ID:20003)

<table>
<thead>
<tr>
<th>Small Molecule</th>
<th>Cell Line</th>
<th>Assay compound conc (uM)</th>
<th>Day</th>
<th>Image</th>
<th>Cell count</th>
<th>Interphase count</th>
<th>Apoptotic count</th>
<th>Dead count</th>
<th>Mitotic count</th>
<th>% Interphase</th>
</tr>
</thead>
<tbody>
<tr>
<td>BI-2536</td>
<td>5637</td>
<td>11.11</td>
<td>1</td>
<td><img src="#" alt="Image1" /></td>
<td>695</td>
<td>543</td>
<td>142</td>
<td>0</td>
<td>10</td>
<td>78.13</td>
</tr>
<tr>
<td>BI-2536</td>
<td>5637</td>
<td>11.11</td>
<td>2</td>
<td><img src="#" alt="Image2" /></td>
<td>438</td>
<td>159</td>
<td>229</td>
<td>1</td>
<td>49</td>
<td>36.30</td>
</tr>
<tr>
<td>BI-2536</td>
<td>5637</td>
<td>11.11</td>
<td>3</td>
<td><img src="#" alt="Image3" /></td>
<td>813</td>
<td>124</td>
<td>641</td>
<td>15</td>
<td>33</td>
<td>15.25</td>
</tr>
<tr>
<td>—</td>
<td>5637</td>
<td></td>
<td>1</td>
<td><img src="#" alt="Image4" /></td>
<td>1207</td>
<td>1168</td>
<td>22</td>
<td>0</td>
<td>17</td>
<td>96.77</td>
</tr>
<tr>
<td>—</td>
<td>5637</td>
<td></td>
<td>2</td>
<td><img src="#" alt="Image5" /></td>
<td>2343</td>
<td>2201</td>
<td>57</td>
<td>2</td>
<td>83</td>
<td>93.94</td>
</tr>
<tr>
<td>—</td>
<td>5637</td>
<td></td>
<td>3</td>
<td><img src="#" alt="Image6" /></td>
<td>4963</td>
<td>4691</td>
<td>252</td>
<td>20</td>
<td>0</td>
<td>94.52</td>
</tr>
<tr>
<td>BI-2536</td>
<td>5637</td>
<td>11.11</td>
<td>1</td>
<td><img src="#" alt="Image7" /></td>
<td>585</td>
<td>457</td>
<td>118</td>
<td>5</td>
<td>5</td>
<td>78.12</td>
</tr>
<tr>
<td>BI-2536</td>
<td>5637</td>
<td>11.11</td>
<td>2</td>
<td><img src="#" alt="Image8" /></td>
<td>448</td>
<td>184</td>
<td>207</td>
<td>5</td>
<td>52</td>
<td>41.07</td>
</tr>
</tbody>
</table>

---

**Download:** Excel (XLSX), Comma-separated (CSV)
OMERO & BIO-FORMATS:
HMS LINCS Database: 27 Channel Cyclic IF

Jia-Ren Lin, Peter Sorger, Harvard Med School
Welcome to the JCB DataViewer! The JCB DataViewer facilitates viewing, analysis, and sharing of multi-dimensional image data associated with articles published in The Journal of Cell Biology.

View

Notochord vacuoles are lysosome-related organelles that function in axis and spine morphogenesis
Kathryn Ellis, Jannera Bagwell, Michel Bagnat

Promoter- and RNA polymerase II–dependent hsp-16 gene association with nuclear pores in Caenorhabditis elegans
Sabine Rohner, (...), Petar Meister

Conserved and divergent features of kinetochores and spindle microtubule ends from five species
J. Richard McIntosh, (...), Ekaterina L. Grischuk
OMERO & BIO-FORMATS: EMDataBank / PDBe

EMD-3216 › Volume slicer

In situ sub-tomogram average of the host-free Chlamydia trachomatis type III secretion system

Sample name: Chlamydia trachomatis type III secretion system (host-free)
Method: Sub-tomogram averaging
Resolution: 33Å (FSC 0.5, gold-standard)

Min: 1.03
Max: 5.04

Quick links
- EMD-3216 overview
- Function and Biology
- Experiments and Validation
- View
- Downloads
- Volume viewer
- Volume slicer
- Visual analysis

Top view
Front view
Right view

Ardan Patwardhan, Ingvar Lagerstadt & Gerard Kleywegt, EBI
MISSION & SCOPE

Biological Imaging

Service
in imaging technologies

Sharing
expertise &
best practice

Innovation
in imaging
technologies

Networking
of imaging facilities
in Europe

Medical Imaging

Access
to imaging
technologies

Data
storage &
analysis

Coordinated
training
programs

http://eurobioimaging.eu
Study metadata and ontologies

Your records and analyses

output

study file
library or assay file
results file

experimental metadata

study type
imaging method
study description
contacts

reagent identifiers
gene identifiers
controls
quality control

measurements
reproducibility
phenotypes

Cellular Phenotype Database
MAGE-TAB
isatab
Image Data Repository (IDR)

Experimental metadata

Integrated studies

Ontological annotations

Thumbnails

Feature vectors

Download local analysis

Cloud analysis

Cross-data browsing
## IDR reality

<table>
<thead>
<tr>
<th></th>
<th>Nov ’15</th>
<th>Jun ‘16</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw image data</td>
<td>37 TB (4 servers)</td>
<td>42 TB</td>
</tr>
<tr>
<td>Total image files</td>
<td>8.2 M</td>
<td>16.5 M</td>
</tr>
<tr>
<td>Total sets</td>
<td>2500</td>
<td>3400</td>
</tr>
<tr>
<td>Total Images</td>
<td>1.3 M</td>
<td>2.8 M</td>
</tr>
<tr>
<td>Total planes</td>
<td>28.9 M</td>
<td>36.1 M</td>
</tr>
</tbody>
</table>

Experimental metadata

Ontological annotations

Feature vectors

http://idr-demo.openmicroscopy.org
Image Data Repository
(http://idr-demo.openmicroscopy.org)
Re-use of IDR data and analytics

```python
import os
import numpy as np
import matplotlib.pyplot as plt
import shelve
import random
from pandas import Series, DataFrame, read_csv
from omero.gateway import BlitzGateway
from matplotlib.widgets import Slider
from matplotlib import gridspec
from sklearn.externals import joblib

from sklearn import random_projection
import sklearn.neighbors as nn
import sklearn.manifold as man
import sklearn.decomposition as dec
from sklearn.preprocessing import scale
from sklearn.metrics.pairwise import

%matplotlib nbagg
plt.rcParams['image.cmap'] = 'gray'
```
PRACTICAL USE CASE....
Common problem?

siControl

siBod1

Iain Porter
<table>
<thead>
<tr>
<th>Time (hours)</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>29</td>
</tr>
<tr>
<td>2</td>
<td>676</td>
</tr>
<tr>
<td>4</td>
<td>1481</td>
</tr>
</tbody>
</table>

Intensity above bkg
Image Quantification and Phenotypes

(a) Controlsi, Controlsi, Bod1si

(b) PP2A-B56α levels

(c) Controlsi prometaphase, Controlsi metaphase, Bod1si, Sgo1si

(d) Quantification of Sgo1 at centromere

(f) 0 min, 25, 75, 95, 120

(g) Control siRNA, Bod1 siRNA, Sgo1 siRNA

Iain Porter 2013
OMERO IN THE MIX
OMERO setup

servername.institution.org

OMERO Server

Lab Group
Lab Group
Facility Group
Public Group
...

Database

OMERO/Institution Login

INTERNET
OMERO group and user system
Groups
Remote Collaboration

- Assessment of abundance of various proteins at the centromere in different mitotic phases
- 50 images from each of 4 time points, 4 combinations of stains, two experimental conditions
- 1600 images

- Acquire
- Store
- Sort
- Analyse
Imaging Infrastructure
Remote Collaboration

- Assessment of abundance of various proteins at the centromere in different mitotic phases
- 50 images from each of 4 time points, 4 combinations of stains, two experimental conditions
- 1600 images

- Acquire ✓
- Store ✓
- Sort □
- Analyse □
Group Rendering and Tagging
○ Flexible viewing of images

○ Filter by:
  - Tagged (general or named tags)
  - Images with/out ROIs
  - Images with comments
Remote Collaboration

- Assessment of abundance of various proteins at the centromere in different mitotic phases
- 50 images from each of 4 time points, 4 combinations of stains, two experimental conditions
- 1600 images

- ✔ Acquire
- ✔ Store
- ✔ Sort
- ☐ Analyse…
Distance Measurements

2.4 μm

15.95°
OMERO.scripts: Channel offset example
ImageJ and OMERO
The Scale of Intensity Analysis

- Andrews et al. 2004: “9x10 pixel box positioned at the plus end of kinetochore fibers. Integrated pixel intensities were measured for over 20 individual fibers in control and Aurora B RNAi cells.”
  - **Duration: > 1 week**

- Porter et al. 2007: 15px diameter circular ROI placed over approx. 10 kinetochores per cell. 10 cells analysed per condition.
  - **Duration: 1 week per experiment**

- Porter et al. 2013: 3 replicates; in each replicate 10 cells (>1,000 kinetochores) measured per condition.
  - **Duration: ~3 hours**
ANALYSIS WITH OMERO
ImageJ and OMERO
Users Can Develop Too!
Some useful links

- OMERO Downloads:
  - http://downloads.openmicroscopy.org/omero/

- OMERO Help Pages:
  - http://help.openmicroscopy.org/

- OMERO Forums:
  - https://www.openmicroscopy.org/community/

- OMERO demo server:
Thank to Funders

HORIZON 2020

MULTI

CORBEL

GLOBAL BIOIMAGING

growing collaboration

EURO-BIOIMAGING

bbsrc

biotechnology and biological sciences
research council

WELLCOME
OME Consortium

Prof Jason Swedlow
University of Dundee

Prof Richard Baldock
MRC Human Genetics Unit, Edinburgh

Prof Rafael Carazo Salas
University of Cambridge

Dr Alvis Brazma
European Bioinformatics Institute, Cambridge

Dr Ilian Davis
University of Oxford

Prof Paul French
Imperial College, London

Dr Spencer Shorte
Institut Pasteur, Paris

Dr Edouard Bertrand
ROI Imaging Montpellier

Dr Gianluigi Zanetti
CRS4, Sardinia

Dr Ilya Goldberg
National Institute on Ageing, Baltimore

Dr Robert Murphy
Carnegie Mellon University, Pittsburgh

Mr Kevin Eicieri
University of Wisconsin, Madison

Prof Gaudenz Danuser
UT Southwestern, Dallas

Dundee, UW Madison, UT Southwestern, Oxford, CRS4, Montpellier, Edinburgh, CMU, Imperial, NIA, Institut Pasteur, EMBL-EBI, Glencoe Software